Neuromuscular Control of Vocal Loudness in Adults as a Function of Cue

by

Andrea Tam

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Speech-Language Pathology

Department of Communication Sciences and Disorders

University of Alberta

Abstract

Background. Speech production in noisy environments is known to elicit the *Lombard effect*, which has been described as an involuntary increase to vocal loudness against background noise. Many studies have explored the auditory-perceptual mechanisms that contribute to the Lombard effect; however, the neuromuscular control mechanisms underlying the physiological adjustments for speech produced during the Lombard effect are not well understood. Together, data from animal, acoustic, and kinematic studies have suggested that control of vocal loudness adjustment during the Lombard effect is systematically distinct from loudness adjustment in response to explicit instruction. The present study examined the effects of different vocal loudness cues on respiratory control, as measured by vocal sound pressure level, lung volume, and chest wall intermuscular coherence during speech breathing tasks in healthy young adults.
Methods & Analysis. Fifteen healthy young adults (20-32 yrs) were recorded during a wordless picture storytelling task, standardized sentence repetition task, and a maximum phonation task in three cue conditions: 1) At conversational loudness, which served as a baseline condition; 2) Verbal instruction to speak at perceived twice conversational loudness, and 3) In multi-talker noise (inducing Lombard changes). The following dependent variables were examined: acoustic recordings of vocal sound pressure level (dB SPL), lung volume events (lung volume initiations, terminations, excursions, and percent rib cage contribution to lung volume excursion), and intercostal-oblique intermuscular coherence. Exploratory analyses were conducted on fundamental frequency, as well as the relative activation of the intercostal and oblique muscles.
Results. The main findings were that different cues for vocal loudness: *i)* Produced similar increases to vocal SPL; *ii)* Did not result in systematic changes to speech breathing parameters; *iii)* Produced comparable intermuscular coherence values within each task in the 15-59 Hz range,

as well as in the 60-110 Hz range, with overall greater mean peak coherence observed in the 15-59 Hz bandwidth, regardless of cue or task. Exploratory analyses suggested that speakers modulated vocal loudness using systematic laryngeal adjustments, as measured by increased F_0 in the loud speaking conditions. **Conclusion.** Different cues had no effect on the control of speech breathing, indicating that the respiratory control circuits involved in vocal loudness change are stable against perturbation in the healthy adult system when the targeted tracheal pressure is controlled for between different cues to increase loudness. The results add to a growing body of literature in which the application of surface EMG and coherence analyses are used as a noninvasive means for understanding speech motor control. Results from this study may provide a platform to further understand speech breathing in individuals with deficits that affect neuromuscular control of the respiratory system.

Preface

This thesis is an original work by Andrea Tam. This research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name: *Neuromuscular control of vocal loudness in children and adults as a function of cue*; Pro00061081; 21 December 2015.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Drs. Carol Boliek and Jacqueline Cummine for their mentorship, positivity, patience and knowledge during this experience. I would also like to thank my committee members, Drs. Benjamin Tucker and William Hodgetts for their expert advice, and Dr. Dave Collins for serving as the external examiner. A massive thank you to the laboratory team members: Alesha Reed for your invaluable support throughout all stages of this process; Mikayla Ho, Neesha Bhat, Grace Crosby, Michael Warren, Danielle Fonesca, Spenser Halfyard, and Kulpreet Cheema for your hard work in data collection and analysis; and Dr. Reyhaneh Bakhtiari for her work on the coherence program. Last but not least, I would like to thank my parents, Mabel and Vincent Tam, for their unconditional support; my grandparents, Lois and Ben Wong, for their constant positivity; and my incredible group of friends for their unending encouragement.

Research reported in this thesis document was supported by the Alexander Graham Bell Canada Graduate Scholarship (Natural Sciences and Engineering Research Council of Canada) and the Walter H. Johns Graduate Fellowship (University of Alberta).

Table of Contents

Introduction 1

The Lombard effect 1

Mechanisms Underlying the Lombard Effect 2

Respiratory Activity During Speech Breathing 4

Respiratory activity during the Lombard effect 6

Coherence 8

Purpose 11

Primary Aims 11

Secondary Aims 12

Hypotheses 12

Methods 13

Participants 13

Procedures 14

Experimental Overview 14

Experimental Tasks & Conditions 15

Equipment 17

Calibration Procedures 20

Sample Size 22

Data Analysis 23

Acoustic Analysis. 23

Chest wall kinematic and lung volume analysis 26

Surface electromyographic and coherence analysis 27

Interrater Reliability 29

Intrarater Reliability 30

Results 31

Effect of cue on vocal SPL 31

Effect of cue on lung volume 33

Effect of cue on chest wall intermuscular coherence 36

Exploratory Analyses 38

Discussion 44

Effect of cue on vocal SPL 44

Effect of cue on lung volume 45

Effect of cue on chest wall intermuscular coherence 49

Contributions of laryngeal mechanisms to vocal loudness adjustment 54

Limitations 55

Future Directions 58

Conclusions 61

References 62

Appendix A 72

List of Tables

Table 1. Estimates of multi-talker noise level derived from a 2cc-coupler (Audioscan Verifit).

Table 2. Interclass correlation coefficients (ICC) and the 95% confidence interval (95% CI) of the ICC values as a measurement of interrater reliability for 10% of the data.

Table 3. *F* ratios and probability values for repeated measures ANOVAs on the effect of cue on mean vocal SPL.

Table 4. Results from post-hoc paired samples t-tests (two-tailed) and probability values comparing the effect of cue on mean vocal SPL for each task.

Table 5. F ratios and probability values for repeated measures ANOVAs on lung volume at utterance initiation (LVI), lung volume at utterance termination (LVT), lung volume excursion used during the breathing task expressed in percent vital capacity (%LVE), and percent rib cage contribution to lung volume change (%RC).

Table 6. Mean and standard deviation effects of cue on lung volume by cue condition and task.

Table 7. F ratios and probability values for repeated measures ANOVAs on significant (above the 95% confidence limit) peak coherence values in two frequency bands: Low (15-59 Hz) and High (60-110 Hz).

Table 8. Mean and standard deviation effects of cue on raw peak coherence values in the low (15-59 Hz) and high (60-110 Hz) bands.

Table 9. Summary of the frequencies (Hz) at which mean peak coherence occurred in the low (15-59 Hz) and high frequency bandwidths (60-110 Hz) for each task and condition across all subjects who demonstrated nonzero coherence (significant at the 95% confidence interval).

Table 10. F ratios and probability values for repeated measures ANOVAs on mean fundamental frequency (F0), intercostal (%MVCIC), and oblique muscle activity (%MVCOB), expressed as a percentage of maximum voluntary contraction.

Table 11. Mean and standard deviation effects of cue on intercostal and oblique muscle activity.

Table 12. Bivariate correlations (Pearson's r) between mean dB SPL and four main outcome variables representing biomechanics of vocal loudness adjustment: Fundamental frequency (F0), lung volume at utterance initiation (LVI), intercostal and oblique muscle activity expressed as a percentage of maximum voluntary contraction (%MVCIC, %MVCOB, respectively).

Table 13. Comparison of findings across various studies investigating the effect of different cues on selected parameters of speech breathing.

List of Figures

Figure 1. An exemplar of a raw acoustic file in Praat for analysis.

Figure 2. Raw waveforms data showing rib cage (RC) and abdominal (AB) kinematic signals, lung volume (LV) corrected for respective RC and AB contributions (see Methods section), intercostal (IC) and oblique (OB) EMG activity, and the acoustic signal during each experimental task for a representative participant.

Figure 3. Coherence spectra tracings during sentence repetition for a representative participant.

Figure 4. Results from post-hoc paired samples t-tests (two-tailed) indicating significant differences in mean vocal SPL between cue conditions for each task.

Figure 5. Results of one-way repeated measures ANOVAs testing for effects of cue on lung volume variables derived from chest wall kinematics.

Figure 6. Results from one-way repeated measures ANOVAs on transformed mean peak coherence values (*z*) for each task in the low (15-59 Hz) and high (60-110 Hz) frequency bands.

Figure 7. Results from paired samples *t*-tests (two-tailed) comparing transformed mean peak coherence values (*z*-scores) between the low and high frequency bandwidths for each task and cue condition.

Figure 8. Results from post-hoc paired samples *t*-tests comparing mean fundamental frequency produced in each task and cue condition.

Figure 9. Results from post hoc paired samples *t*-tests comparing transformed mean peak coherence values (*z*) produced during resting breathing (non-speech task) and storytelling in the CONV condition (speech task) in the low frequency band (15-59 Hz) and the high frequency band (60-110 Hz).

Figure 10. Mean peak intercostal-oblique intermuscular coherence values (*r*) for each subject (S1-15) in the low and high frequency bands, showing individual responsiveness to each cue condition across all tasks.

Introduction

The Lombard effect

Speech motor control refers to the mechanisms that mediate formulation of the linguistic (propositional) message and the subsequent acoustic speech signal. In order for people to speak, the brain must plan, program, control and execute precise and well-timed movement commands to the respiratory, laryngeal, velopharyngeal, and oral-articulatory subsystems of the speech mechanism (Duffy, 2013). These central commands result in contraction of over 50 paired muscles to produce speech (Perkell, 2012). Importantly, the centrally-driven commands to the peripheral speech mechanism are adaptive; they can be varied to allow us to effectively speak in a variety of communicative settings.

The *Lombard* effect is one well-known example of the way speech production is adapted in response to the environment. First described by French otolaryngologist Étienne Lombard in 1911, it is understood as an automatic, unconscious increase to vocal sound pressure level (SPL) in the presence of sufficient background noise (Brumm & Zollinger, 2011). Since Lombard's initial discovery, the effect has been documented at noise levels of about 80 dB SPL or higher (e.g. Pick, Siegel, Fox, Garber & Kearny, 1989; Summers, Pisoni, Bernacki, Pedlow & Stokes, 1988), but also has been observed at relatively lower noise levels around 55 to 60 dB SPL (e.g. Winkworth & Davis, 1997; Tufts & Frank, 2003). Compared with speech produced in quiet conditions, vocal SPL is reported to increase anywhere between 2 dB (Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988) to 15.7 dB (Patel & Schell, 2008), with the magnitude of change generally increasing with background noise level. In addition to increased vocal SPL, a suite of acoustic changes has been documented in the presence of noise, including flattening of spectral slope (Cooke & Lu, 2010), as well as increased word duration and elevation of fundamental frequency (Patel & Schell, 2008; Van Summers, Pisoni, Bernacki, Pedlow & Stokes, 1988).

Mechanisms Underlying the Lombard Effect

Two primary processes have been proposed to underlie the Lombard effect. First is the detection of mismatches between the targeted and produced output by the *auditory-perceptual system*. Speech in background noise is thought to represent such a mismatch because noise presumably degrades or alters self-produced acoustic output relative to the expected production (Eliades & Wang, 2012). Studies of noise-induced vocalizations in decerebrate cats (Nonaka, Takahashi, Enomoto, Katada, and Unno, 1997) and squirrel monkeys (Hage, Jürgens, Ehret, 2006) have been used to contend that brainstem feedback mechanisms are fundamental to detection and initiation of noise-induced shifts in vocal SPL. However, single neuron electrode recordings from the *marmoset* auditory cortex indicate that higher-level mechanisms also may be involved during the Lombard effect. Eliades and Wang (2012) found that when marmoset monkeys vocalized in white noise, the resulting increase to vocal SPL was accompanied by a systematic shift in the firing pattern of auditory neurons toward activity levels observed under quiet conditions, thus acting as a putative error-detection mechanism. Although neural recordings have not been carried out in humans, there are magnetoencephalographic findings that also show increased activity in the human auditory cortex when speech is produced in the presence of background white noise relative to quiet speaking conditions (Houde, Nagarajan, Sekihara, & Merzenich, 2002).

Second, it is assumed that detection of the mismatches between the targeted and actual acoustic output triggers corrective *physiological adjustments* to the speech system to repair the speech signal and allow the speaker to maintain communicative effectiveness in noise (Lane &

Tranel, 1971). That is, an integrated speech perception-speech production system drives the Lombard effect. Insights into how this system operates during the Lombard effect have largely been extrapolated from acoustic data. Noise-induced increases to vocal SPL persist even when speakers are informed about the Lombard effect and are instructed to suppress it, thus supporting its description as an automatic feedback response that cannot be effectively inhibited (Pick Jr., Siegel, Fox, Garber, Kearney, 1989). However, the effect can be modified to some extent because both the semantic content of an utterance (Patel & Schell, 2008) and communicative intent of a task (Garnier, Henrich, & Dubois, 2010) have been shown to increase the magnitude of vocal SPL adjustment in noise. Presumably, these acoustic changes promote communicative effectiveness; indeed, speech perception literature suggests that increases to vocal SPL under the Lombard effect improve the signal-to-noise ratio of the acoustic speech signal (Cooke & Lu, 2010) and enhance the intelligibility of the speaker's message in noise (Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Lu & Cooke, 2009). Taken together, it appears that external noise influences speech production in complex ways, and that there may be higher order perceptual and cognitive mechanisms at work during this relatively automatic adjustment to vocal loudness.

Whereas the acoustic changes that accompany noise-dependent increases to vocal SPL have been well described (e.g. Lane & Tranel, 1971; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Patel & Schell, 2008; Lu & Cooke, 2009; Cooke & Lu, 2010; Stowe & Golob, 2013), studies of the speech mechanism are needed in order to provide a more direct window into the motor control systems and strategies used to regulate vocal SPL during the Lombard effect. Adult speakers use both the respiratory and laryngeal subsystems to increase vocal SPL; however, previous work has shown that respiratory mechanisms predominate increased pressure generation for loud speech (Stathopoulos & Sapienza, 1993; Stathopoulos & Sapienza, 1997). Therefore, this study will focus on respiratory mechanics used during adjustments to vocal SPL.

Respiratory Activity During Speech Breathing

The respiratory system, which is comprised of the pulmonary apparatus (trachea, pulmonary airways, and lungs) and chest wall (rib cage, diaphragm, abdomen, and abdominal contents) (Hixon, Mead, & Goldman, 1976), drives the variations in air pressures and flows that are converted to acoustic energy in the speech signal. During speech production, the respiratory system performs differently than during resting (tidal) breathing. Whereas tidal breathing is driven by the need for gas exchange and is associated with relatively regular phases of inspiration and expiration, breathing behaviours during speech (speech breathing) are varied in order to efficiently regulate various aspects of spoken communication. Four parameters can be used to describe the relationship between biomechanical behaviours of respiratory system and perceptual characteristics of speech breathing during a given task: Lung volume, tracheal pressure, shape, and flow are linked to the perceptions of breath group length, vocal loudness, inspiratory duration, and voice quality, respectively (Hixon, 1993; Hoit & Hixon, 1987). Flow of air from the lungs is highly dependent on valving and resistance created by laryngeal and oral structures; as such, flow will not be discussed for the purposes of this review.

Control and coordination of the respiratory system is closely linked to lung volume because lung volume status dictates whether passive (i.e. elastic recoil of the lungs and rib cage) or active muscular forces (i.e. muscular contraction) predominate during speech. The amount of lung volume used for a given breathing task can be expressed in relation to *vital capacity* (VC), which refers to the maximum volume of air that can be expired after maximum inspiration. To illustrate the relationship between lung volume and muscular force, consider that when a person is asked to take a deep breath and sustain production of a vowel (e.g. /ɑ/) for as long and as steady as they can (a maximum phonation task), they are using their full range of VC: Phonation begins at high lung volumes (the top of vital VC), where passive forces cause the rib cage and lungs to recoil inward, thus forcing air from the lungs. As phonation progresses across time, lung volume decreases toward end expiratory level and muscular force is increasingly applied in the expiratory direction in order to maintain steady and adequate tracheal pressure (Hixon, 1973; Hixon, Mead, & Goldman, 1976; Hixon & Weismer, 1995).

In contrast, conversational level speech is typically produced within the midrange of VC, beginning at lung volumes between 50 to 60 %VC and ending at approximately 50 to 30 %VC, with an average lung volume excursion of 20 %VC (Hixon, Mead, & Goldman, 1976). It is biomechanically efficient to produce speech in this range because passive recoil pressures are close to those required for the targeted tracheal pressure, thus requiring minimal active expiratory muscular effort (Hixon, Goldman, & Mead, 1973). In adults, approximately 5-8 cmH_2O (Smitheran & Hixon, 1981) are required to produce and sustain speech at conversational loudness. For speech that is louder than that used during conversation, tracheal pressures of approximately 10-20 cmH_2O have been reported (Stathopoulos & Sapienza, 1993; Ladefoged & McKinney, 1963). Accordingly, higher lung volumes at utterance initiation, increased muscular activity of the chest wall, and greater proportions of overall lung volume have been documented at higher tracheal pressures (Stathopoulos & Sapienza, 1993, Stathopoulos & Sapienza, 1997; Dromey & Ramig, 1998; Winkworth, Davis, Ellis, & Adams, 1994)

Although the rib cage and abdomen move continuously to effect lung volume change during speech, adults assume a background chest wall shape where the abdomen is displaced inward (smaller than at rest) and the rib cage is displaced outward (larger than at rest), which serves to expand the rib cage and lift and dome the diaphragm (Hixon, Goldman, & Mead, 1976). The mechanical advantages of this chest wall shape for vocal loudness modulation are two-fold: First, the expiratory rib cage and abdominal muscles are at a prime length-tension characteristic for generating movements that result in vocal loudness change. Second, inward displacement of the abdominal wall promotes efficient expiratory rib cage movement. This is integral to effective vocal loudness modulation because the contribution of the rib cage to lung volume change has been shown to increase with tracheal pressure (Stathopoulos & Sapienza, 1997).

Respiratory activity during the Lombard effect

Thus far, speech breathing parameters have been described in terms of the strategies employed during voluntary adjustments to loudness, such as when a subject is instructed to speak above their usual conversational vocal SPL. There are data to suggest, however, that the respiratory patterns employed during vocal loudness adjustment can vary according to *cue*; that is, whether a speaker increases SPL intentionally in responses to instruction, or does so involuntarily against background noise, as in Lombard effect. One study demonstrated that, despite a uniform 10 dB SPL increase to vocal loudness, healthy young adults initiated utterances at higher lung volumes, terminated utterances at higher abdominal volumes, and used a larger range of overall lung volume when reading sentences in noise than when instructed to read the same sentences at perceived twice conversational loudness (Huber, Chandrasekaran, & Wolstencroft, 2005). Similar patterns have been observed during monologue tasks, in which speakers use comparatively higher lung volume initiations and greater excursions during loud speech cued by noise versus verbal instruction (Huber, 2007). From these findings, it appears that physiological adjustments of the chest wall during the Lombard effect are more efficient than those used in response to explicit instruction. Specifically, relatively higher lung volume initiations suggest distinct reliance on passive recoil mechanisms during the Lombard effect, thus reducing the muscular work needed to generate the higher targeted tracheal pressures for loud speech.

Clinical data offer some support to this argument. Stathopoulos et al. (2014) found that adults with Parkinson's disease, who have difficulties speaking loud enough to be heard during conversation, read passages and produced monologues at greater SPLs with improved tracheal pressure, vocal fold adduction, and higher lung volume initiations in the presence of background multi-talker babble relative to quiet conditions. Therefore, it appears that the Lombard effect promotes increased biomechanical efficiency of the speech mechanism in both healthy speakers as well as those who have neurologically compromised systems.

Although these findings are intriguing, the studies that explored monologue production found that vocal SPL was systematically greater when loud speech was produced in response to background noise than verbal instruction (Huber, 2007; Stathopoulos et al. 2014). As a result, it is unclear whether the Lombard effect truly elicits distinct speech breathing patterns because the effects of vocal SPL could have potentially driven the observed differences in respiratory kinematic patterns. Therefore, it is crucial that any replication and extension of this work control for potential differences in vocal SPL adjustment between loud speaking conditions. Further, claims regarding reduced muscle effort during the Lombard effect can only be inferred from these kinematic data. As such, employing techniques that can directly observe muscular activity, like coherence analysis, in conjunction with kinematic analysis can provide a more complete description of the neuromuscular control and coordination mechanisms underlying the Lombard effect.

Coherence

In the larger motor control literature, *coherence analysis* has been used as a non-invasive means of assessing physiological discharges to muscle, with the larger goal of characterizing functional neuromuscular control networks. Coherence analysis is based on squared cross-correlation of oscillations from two separate physiological signals in the frequency domain (Grosse, Cassidy, & Brown, 2002). For example, coherence measurements can be derived from pairing electroencephalographic (EEG) or magnetoencephalographic (MEG) with electromyographic (EMG) recordings (*corticomuscular coherence*) or from multiple EMGs over paired muscle areas (*intermuscular coherence*). Coherence values are bound between 0 and 1 and indicate the strength of correlated oscillatory activity between different signal sources at a particular frequency: A value of 1 indicates a perfect linear relationship between signals, whereas a value of 0 indicates no linear relationship. Two signals are therefore said to be coherent when they exhibit a significant frequency-specific linear relationship. Because intermuscular coherence analysis is derived from paired EMG signals, its use is particularly appropriate for investigations that aim to characterize the underlying mechanisms driving the coordination of disparate muscle areas (Boonstra, 2013).

Frequency of intermuscular coherence. Although the precise mechanisms underlying intermuscular coherence continue to be discussed (Boonstra, 2013), findings from typical (Maurer, von Tscharner, & Nigg, 2013; Jaiser, Baker, & Baker, 2016) and clinical populations (Hansen et al., 2005; Norton & Gorassini, 2006; Fisher, Zaaimi, Williams, Baker, & Baker, 2013) have contributed to a general consensus that intermuscular coherence represents common descending oscillatory drive to disparate muscle areas during coordinated muscle activity, with output likely originating, in part, from motor cortex (Grosse, Cassidy, & Brown, 2002). In studies of coherence between muscles of the limbs, synchronized oscillations have been primarily observed in two primary bandwidths: the β-band (15-30 Hz) and γ-band (30-60 Hz) (Maurer, von Tscharner, & Nigg, 2013; Jaiser, Baker, & Baker, 2016). Much of the available data on intermuscular coherence are based on studies of limb movement; however, coherent oscillations in the β- and γ-bands also have been reported between muscles of the speech mechanism, including strap muscles of the neck (Stepp, Hillman, & Heaton, 2011), muscles of the jaw (Smith & Denny, 1990), and respiratory muscles of the chest wall (Smith & Denny, 1990; Denny & Smith, 2000; Tomczak, Greidanus, & Boliek, 2013) during nonspeech (e.g., chewing, controlled deep breathing), speech (e.g., reading aloud), and speech-like tasks (e.g., reading silently). In addition, higher frequency coherence between 60-110 Hz has been found in the chest wall muscles (Bruce & Ackerson, 1986; Smith & Denny, 1990).

It is interesting to note that the coordination requirements of a given task appear to modulate the frequency bandwidth at which coherence occurs. Smith and Denny (1990) found significant 60-110 Hz chest wall coherence during a deep breathing task, in which subjects inhaled and exhaled according to a visually-presented waveform. When subjects read a passage aloud, however, coherence in this range was significantly reduced. These results were replicated in a following study that compared 60-110 Hz coherence during spontaneous monologue production and controlled deep breathing (Denny and Smith, 2000). This pattern, in which high frequency coherence was reduced during speech, was interpreted as a shift in the neuromuscular control of the respiratory system in order to meet the varied demands of breathing during speech production; that is, speech and nonspeech breathing tasks were independently controlled by oscillations of different frequencies (Smith & Denny 1990; Denny & Smith, 2000). Because high frequency oscillations appear to be unique to the respiratory system and are primarily produced during nonspeech breathing, these results have been taken as evidence of brainstem central pattern generator control of the respiratory system in humans (Bruce & Ackerson, 1986; Smith & Denny, 1990).

To date, no work has been reported on whether frequency-specific modulation of intermuscular coherence occurs during speech tasks that potentially rely on distinct underlying coordination patterns of the respiratory system. Since the available animal (Nonaka, Takahashi, Enomoto, Katada, and Unno, 1997; Hage, Jürgens, & Ehret, 2006; Eliades & Wang, 2012) and kinematic data (Huber, Chandrasekaran, & Wolstencroft, 2005; Huber, 2007) suggest that the motor control and coordination mechanisms engaged during the Lombard effect are physiologically distinct from those used when vocal SPL is increased deliberately, it is possible that measurements of chest wall intermuscular coherence have the potential to distinguish the neuromuscular control mechanisms arising from different vocal loudness cues. More specifically, it is possible that noise-related adjustments to vocal SPL might yield coherence in the 60-110 Hz range because of the presumed involvement of brainstem mechanisms during the Lombard effect (Nonaka, Takahashi, Enomoto, Katada, and Unno, 1997; Hage, Jürgens, & Ehret, 2006).

Strength of intermuscular coherence. Task related factors also have been shown to influence the *strength* of coherent activity between separate muscle groups. Tasks of greater difficulty have been shown to reduce β-band coherence; for example, β-band coherence is reduced between the gastrocnemius muscles during fast running relative to slower running (Maurer, von Tscharner, & Nigg, 2013) and between the strap muscles of the neck during a divided attention task (counting backwards aloud) relative to regular reading (Stepp, Hillman, & Heaton, 2010). As suggested by Huber, Chandrasekaran, and Wolstencroft (2005), adjusting vocal SPL in response to verbal instruction may be more difficult, as this task requires greater attention to, and precision control of, an internally calibrated loudness target compared to the Lombard effect, which is thought to operate within an automatic feedback loop. In the case of vocal loudness adjustment then, intermuscular coherence may be sensitive to differences in difficulty between vocal loudness cues.

Reduced β-band coherence also has been reported for maximum capacity tasks that demand use of the entire lung volume excursion range (e.g., maximum phonation, vital capacity maneuvres) than those requiring comparatively less lung volume excursion (e.g., conversational speech, tidal breathing) (Tomczak, Greidanus, & Boliek, 2013). In turn, the magnitude of coherence may offer insight into how chest wall neuromuscular control varies with different vocal loudness cues and the associated demands on lung volume. As discussed previously, relatively larger lung volume excursions observed during the Lombard effect compared with those observed during voluntary vocal loudness adjustments could be related to the higher vocal SPLs and accompanying differences in muscular effort. Therefore, there is a need to examine coherence between chest wall muscles across different loudness cues while also holding changes to vocal SPL constant to determine the extent to which modulation of the respiratory system, under conditions of motor equivalence, uses different neuromuscular control mechanisms.

Purpose

Primary Aims

The primary purpose of the present study was to examine whether or not intermuscular coherence between chest wall muscles in healthy young adults varied in response to different cues to increase vocal loudness. During a series of speech and nonspeech breathing tasks, adults were asked to phonate and speak at conversational loudness (baseline); at a level that was perceived to be twice conversational loudness (voluntary adjustment); and in the presence of background multi-talker babble (involuntary adjustment) that simulated noise commonly encountered in everyday communicative environments, such as when conversing in busy restaurants or on a crowded subway platform.

Secondary Aims

Secondary aims of this study were to replicate previous findings on the acoustic and lung volume changes associated with different vocal loudness cues.

Hypotheses

The specific questions we asked in this study were as follows: First, we explored whether or not chest wall intermuscular coherence produced during the Lombard effect (involuntary adjustment) would be present in the low frequency band (15-59 Hz, comprising the β- and low γ-band) and/or in the high frequency (60 - 110Hz) band and further; whether or not chest wall intermuscular coherence produced in response to explicit instruction (voluntary adjustment) would be primarily reflected in the β band. We hypothesized that *loud speech cued by noise (the Lombard effect) would elicit significant intermuscular coherence in the lower frequency and higher frequency bandwidths, whereas loud speech cued by verbal instruction would produce significant coherence in the lower frequency bandwidth only.*

Second, we asked whether or not task demands would influence the strength of chest wall intermuscular coherence during vocal loudness adjustment. We hypothesized that *loud speech cued by instruction would show reduced intermuscular coherence as compared to loud speech cued by noise (the Lombard effect).*

Methods

Participants

Seventeen healthy young adults were recruited for the study. Criteria for participant selection included: a) Typical speech and language as determined by self-report; b) Normal hearing as determined by a pure-tone audiometric screening at 25 dB HL at 1000, 2000, and 4000 Hz bilaterally (ASHA, 1997; Alberta College of Speech-Language Pathologists and Audiologists, 2015); c) No history of respiratory disease or neurological disease as determined by self-report; d) No history of an acute or chronic skeletal or muscle condition affecting the abdominal or thoracic regions as determined by self-report; and f) No history of smoking in the past year. Two participants were excluded from the study due to equipment malfunction ($n = 1$) and failure to pass audiometric screening ($n = 1$). Therefore, a total of fifteen participants met inclusion criteria and participated in this study ($M = 24.5$, $SD = 3.5$; age range: 20-32; sex: 12 women, 3 men; body mass index: $M = 26.09$, $SD = 3.89$).

Prior to participating in this study, all participants provided written and informed consent. The University of Alberta Health Research Ethics board approved this study and associated experimental procedures.

Procedures

Experimental Overview

All testing took place over one session. Following written and informed consent, participants were familiarized with the study protocol and equipment. Next, participants underwent a hearing screening as described above using a calibrated pure tone audiometer (Maico MA-25) in a quiet room. Once the screening was completed, results were shared with the participant and a paper copy of the results was kept for reference. In the event that a participant failed the hearing screening (i.e., failed to respond to 2 out of 3 presentations at any frequency in one or both ears), the participant had their hearing re-screened in a sound attenuated booth. Participants were excluded from the study if this re-screening resulted in failure and were advised to speak with their physician regarding a referral to a registered Audiologist for further testing.

Acoustic recordings, chest wall kinematics, and EMG signals were collected from all participants during speech tasks (storytelling; sentence repetition) and nonspeech tasks, including resting tidal breathing and maximum phonation. These tasks provided the opportunity to examine coherence between chest wall muscles across a range of lung volumes, tracheal pressures, and in the presence or absence of cognitive-linguistic demand.

Participants were asked to maintain an upright, seated posture with their backs unsupported during all calibration and experimental procedures. During experimental tasks, participants sat directly facing the investigator at a distance of 60-80 cm measured from the mouth of the speaker to that of the investigator. This distance remained constant within each session in order to control for potential variability of responses driven by proximity1. A range in distance was applied to accommodate height differences between participants and the investigator. One to two research assistants were present during each session in order to assist with data acquisition. For all conditions, maximum phonation occurred first and storytelling occurred last. Participants were invited to take short breaks and to drink water throughout the study to limit fatigue. Sessions lasted approximately 1 to 1.5 hours.

Experimental Tasks & Conditions

Participants completed experimental tasks in the following cueing conditions: 1) CONV: This condition served as baseline vocal SPL. Participants were asked to speak (storytelling task and sentence repetition task) or phonate (maximum phonation task) at their perceived conversational loudness ("Say this/tell me this story using your typical conversational voice"); 2) CUED: Participants were asked to speak or phonate at a level perceived to be twice conversational loudness ("Now I want you to speak at what you feel is twice as loud as your typical conversational voice"); 3) NOISE: Participant listened to multi-talker noise presented binaurally through insert earphones at a level that was set to elicit a +7 dB SPL change from conversational SPL. Participants were told that they would be listening to noise that "sounded like a lot of people talking at the same time", but received no instructions as to how loud to phonate or talk, as this cue served as a naturalistic manipulation of vocal SPL. Participants were always exposed to the CONV condition first followed by the either the CUED or NOISE conditions, which were counterbalanced.

---

1 Bremmekamp & Boliek (2015, unpublished) utilized a procedure in which they walked away from participants in order to visually cue a 5-7 dB SPL increase in vocal level. During piloting for the present study, providing a visual cue (i.e., walking backwards) appeared to effectively increase vocal SPL.

Simultaneous acoustic, chest wall kinematic, and EMG recordings were collected from participants during three experimental tasks: 1) Phonation: A minimum of nine trials of a maximum phonation task (three during the CONV condition; three during the CUED condition; three during the NOISE condition), in which participants were instructed to take a deep breath and say “ah” for as long and as steady as they could; 2) Sentence repetition: Twenty-seven trials of a sentence repetition task (nine in the CONV condition; nine in the CUED condition; nine in the NOISE condition), in which participants repeated sentences produced by the investigator, in the following order: “Buy Bobby a puppy”, “the blue spot is on the key”, and “the potato stew is in the pot”. These sentences were chosen because they represent all vertices of the vowel quadrilateral and would be most comparable to the storytelling task, where participants would likely produce a variety of vowel shapes. The investigator presented sentences at conversational level across all conditions. Prior to the NOISE condition, the investigator reminded the participants of the stimulus sentences and had participants practice the sentences in quiet. 3) Storytelling: Three trials of storytelling participants (one in the CONV condition; one in the CUED condition; one in the NOISE condition) using the Edmonton Narrative Norms Instrument (ENNI, Schneider, Dubé, & Hayward, 2005). Stories A1, A2, B1, and B2 were included in the study as they were considered to elicit speech samples of adequate length based on pilot data. Stories were randomized and assigned to each cueing condition. Each story was printed single sided and placed in its own duotang. Participants were instructed that they would be asked to tell a story about the pictures they saw. During task administration, the investigator held the duotang and flipped pages to ensure participants remained seated upright and to minimize extraneous movements of the trunk and chest wall.

Equipment

Acoustic Recordings. Acoustic recordings were used to obtain measures of vocal sound pressure level (dB SPL) during speech and phonation tasks. Vocal sound pressure levels (SPL) were digitally recorded using a Shure MX-185 directional microphone, which was placed at the forehead midline, 10 cm away from the mouth opening. Signals were amplified (M-Audiobuddy Pre-Amplifier) and recorded at a sampling rate of 44.1 kHz on a laptop computer using TF32 Software (Milenkovic, 2001). The same microphone was used to obtain recordings across all sessions. At the end of each session, the microphone was calibrated against a 440 Hz reference tone presented 10 cm from the mouth opening (KORG orchestral Tuner, OT-12). A sound level meter (ExTech sound level meter, 407764) was placed in line with the forehead-mounted microphone and a dB SPL value was recorded once the output readings appeared stable. This value was later applied during acoustic analysis to reflect the actual dB SPL produced during the experimental tasks and conditions.

Chest wall kinematic and lung volume recordings. Variable inductance plethysmography (Respitrace, Ambulatory Monitoring Inc., NY) was used for the acquisition of chest wall kinematics. Respitrace transduction bands were placed around the rib cage and abdomen according to the protocol outlined in Boliek, Hixon, Watson, and Jones (2009). Outputs from each of the transduction bands and the pneumotachometer were simultaneously displayed on a four-channel storage oscilloscope (Tektronix TDS 2014) and recorded on an FM data recorder (DV8, Vetron Technology) to a digital video cassette (AY-DV124AMQ, Panasonic) as separate, but time-locked signals. When an FM data recorder is used in combination with a video camera, all data formats, including the audio waveform can be digitized, recorded, and synchronously played back in their original format for subsequent analyses.

Surface electromyographic recordings. Surface electromyographic (EMG) recordings were extracted to assess intermuscular coherence between the rib cage and abdomen during each experimental task and condition. For each participant, paired electrodes (Kendal Soft-E H69P, Tyco Healthcare Group, Mansfield, MA) were placed on the right side of the chest wall, over the 6^{th} intercostal space and oblique muscle regions, 2 cm apart (center-to-center) and oriented parallel to fiber direction. Intercostal electrodes were placed ventrally 8-10 cm from midline, and the oblique electrodes were placed at a midpoint between the anterior superior iliac spine and caudal border of the rib cage. Following the protocol used by Tomczak, et al. (2013), this electrode placement configuration avoids the intercartilaginous region of the rib cage and enhances ventral-dorsal EMG placement. To increase signal-to-noise ratio and mitigate power line interference occurring at 50 or 60 Hz (Stepp, 2012), a third electrode was placed on the clavicle to serve as a reference signal, or ground. Signals may have reflected some degree of inspiratory muscle activity during expiratory speech and nonspeech breathing tasks given that muscles of inspiration lie superficial to the expiratory internal intercostal muscles. All EMG signals were amplified (Grass P511, Quincy, MA), band-pass filtered (3-3000 Hz), and sampled at 10 kHz. Signals were collected using a multichannel data acquisition system (PowerLab 16SP ML795; ADInstruments, Colorado Springs, CO) and subsequently saved to a computer using LabChart software (v5.5.6; ADInstruments).

Video recordings. Video recordings (Canon ZR60 Camcorder) were collected during the entire study protocol in order to ensure that offline analyses only included data that were free from extraneous limb and trunk movements. A second microphone was attached to the participants' clothing and acoustic signals were digitized (10 kHz) along with the video signals to ensure that an acoustic record of the study was available during analyses.

Multi-talker noise delivery system used to elicit the Lombard effect. Multi-talker noise (Hearing Aid Research Laboratory Speech Intelligibility Tests: *Multitalker Babble*, 1994) was presented binaurally to each participant through ER-2^{TM} insert earphones (Etymotic Research Inc.) via a tablet computer (iPad Air 2, Apple Inc.). Disposable foam ear tips were used with the insert earphones and were changed prior to each new participant. The multi-talker noise audio file was edited into a 15-minute looped recording using GarageBand (Apple Inc., 2012) in order to ensure an uninterrupted noise stimulus was delivered to the participant.

The noise output from the iPad and earphones was calibrated in a 2-cc coupler set to measure live speech (Audioscan Verifit) in order to estimate the level of noise delivered to participants' ears, as well as to estimate the minimum and maximum level of sound output from the tablet. A 2-cc coupler simulates the dimensions of the average adult ear canal and is widely used by hearing aid manufacturers to estimate the output level of hearing aids (Humes & Bess, 2014).

Coupler-based estimates of noise level were defined as the mean of two sound level measurements for each unit of volume level increase using the volume rocker on the iPad. Table 1 provides estimates of the sound level of multitalker noise delivered via insert earphones with each unit of increase in volume on the iPad volume rocker switch.

Number of clicks on iPad volume rocker	Sound level estimate (dB SPL)
1	50.75
2	54.5
3	60
4	65
5	70
6	75
7	79.25
8	83.5
9	87.25
10	91.25
11	95.75
12	99
13	102.5
14	106
15	109.75
16	113.5

Table 1: Estimates of multi-talker noise level derived from a 2cc-coupler (Audioscan Verifit).

Calibration Procedures

Chest wall kinematic and lung volume recordings. A combination of respiratory maneuvers was used to allow conversion of chest wall kinematics into lung volume events. Calibration procedures included: 1) Three trials of a vital capacity maneuver performed in accordance with the American Thoracic Society (Wanger et al., 2005); 2) Three trials of an isovolume maneuver, in which participants sighed out, pulled their abdomen inward, and then allowed it to relax outward while breath holding. Because the rib cage and abdomen provide independent, yet disproportionate contributions to lung volume change (Konno & Mead, 1967), isovolume maneuvers provide a means to calculate a correction factor so the respective signals can be adjusted to reflect an equal and opposite relationship (i.e., a slope of -1) during post hoc analyses; 3) One minute of tidal breathing (approximately l0 tidal breaths), in which participants were asked to breathe into a disposable mouthpiece/bidirectional pneumotachometer system (Hans Rudolph Inc.) while wearing nose clips in order to capture oral airflow. Airflow signals were transduced (Model DP45-14, Validyne Engineering Corp., Northridge, CA) and recorded on a separate channel of the FM data recorder. Lung volume measurements were derived from the time-integrated oral airflow waveforms for each participant and were then compared against the volume displaced by a 3L syringe; 4) Twenty stable tidal breaths were collected for each participant to serve as baseline lung volume.

Surface electromyographic recordings. Participants completed maximum voluntary isometric contraction (MVIC) tasks so that muscular effort could be normalized across participants. Normalization procedures are required to account for individual differences in subcutaneous body fat and skin fold thickness, which are known to effectively low pass filter and attenuate signal amplitude (Lehmann & McGill, 1999). MVIC tasks allow signal amplitudes to be expressed in relation to those obtained during maximal contraction of the muscles of interest, thereby improving the comparability between EMG amplitude measurements (Stepp, 2012; Clair-Auger, Gan, Norton, & Boliek, 2015). The expiratory limbs of the vital capacity maneuvers noted in the above section were used to derive maximum intercostal muscle activity. Three trials of trunk rotation against resistance provided at the opposite (left) shoulder were used to elicit maximum oblique muscle activity. This exercise was sustained for 10 seconds across three trials. The investigator coached participants during each task and provided verbal encouragement throughout.

Multi-talker noise. We anticipated that variability in SPL across loud speaking conditions and any accompanying differences in muscular effort could result in challenging comparisons. Therefore, we aimed to increase participants' vocal SPL by 7 dB above mean conversational SPL during the NOISE condition, because previous data indicate that adults increase vocal SPL by 7 to 10 dB when instructed to speak at levels perceived to be twice as loud as their typical conversational level (Dromey & Ramig, 1998; Huber, Chandrasekaran, &

Wolstencroft, 2005; Tam, Cummine, Hodgetts, Tucker, & Boliek, 2016, unpublished). The following protocol was used to establish target vocal SPL during the NOISE: A microphone (ExTech Microphone Extension Cable, 407764-EXT) attached to a sound level meter (ExTech Sound level meter, 407764) was clipped into the boom of a tripod microphone stand and was positioned in line with participants' mouths at a distance of 30cm. To obtain a coarse but efficient estimate of mean conversational SPL, a research assistant manually recorded all output values from the sound level meter during storytelling in the CONV condition. The two highest and lowest values were removed in order to limit influence from outliers and an average of the next highest and lowest values were obtained. This mean value +7 dB SPL was then recorded as the *target vocal SPL* for the NOISE condition.

Prior to administering the NOISE condition, participants listened to multi-talker noise through insert earphones while producing a 1 to 2 minute monologue about their favourite vacation. During the monologue, the level of multi-talker noise was manually increased using the iPad volume rocker until at least 10 measurements from the sound level meter readout reached the target vocal SPL. This level of noise was held constant during storytelling in the NOISE condition to elicit the Lombard effect at a level that was +7 dB SPL above conversational loudness. In total, participants were exposed to approximately 10 minutes of multi-talker noise during the study.

Sample Size

Sample size was determined by previous work (Mager, Boliek at al., 2015, in preparation), which reported significant within-group differences and a large effect size (d = 1.29) when 12-60 Hz chest wall coherence was compared between two loudness conditions (comfortable loudness and twice conversational loudness) among nine subjects.

Data Analysis

Acoustic Analysis.

To facilitate comparison of dB SPL across conditions, only vowels were included in the analysis. Although vocalic, approximants and rhotic vowels were excluded from analyses on the grounds of variability in forms during casual speech, where visual cues were less reliable (e.g., due to reduction and/or coarticulation with adjacent vowels). Similarly, vowels adjacent to approximants were excluded if boundaries could not be reliably distinguished with visual cues from the spectrogram or waveform. As seen in Figure 1, vowel onset boundaries were marked where formants were visible, ½ cycle to the right of the periodic signal. Vowel offset boundaries were marked at a ½ to the right of the aperiodic signal. However, if glottal pulses were observed at vowel onset or offset (i.e., creaky vowel quality) they were included within segment boundaries, given that glottal pulses could not be consistently or reliably excluded between speakers. Segments were excluded if they were contaminated by electronic noise or were produced while laughing or coughing. Repetitions, pauses, filler words (e.g., "um" or "uh"), and hesitations where participants did not finish their utterance were excluded on the premise that they may reflect cognitive-linguistic planning problems (Jurafsky, Bell, Fosler-Lussier, Girand, & Raymond, 1998) and may therefore be inherently different from utterances that were produced without error.

Audio recordings from each participant were converted into a standard Resource Interchange File Format (.RIFF). Following file conversion, textgrids for each audio recording were manually generated in Praat (v6.0.19, Boersma & Weenink, 2016). To maximize consistency across tasks and dependent measures, recordings were first segmented on a broad level by breath group. Breath group locations were confirmed using the kinematic and coherence data. As seen below in Figure 1, the breath group tier was labelled with the breath group number and an orthographic transcription of speech produced within that breath group. Individual vowels within each breath group were then segmented, transcribed phonemically, and labelled, on a separate *phon* tier using a modified version of ARPAbet, which is a phonemic code that uses the Latin alphabet and does not require use of symbols from the International Phonetic Alphabet Vowel for transcription.

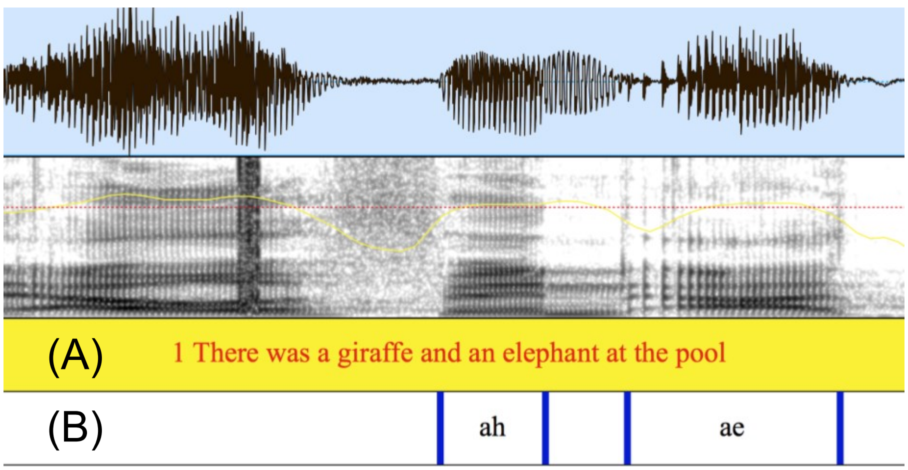


Fig 1. An exemplar of a raw acoustic file in Praat for analysis. (A) denotes the breath group tier (*BG)*, (B) indicates the segmental/vowel tier (*phon*). Modified ARPAbet was used for phonemic vowel transcription. Vocal SPL values were extracted from segments in the *phon* tier (A). Approximants as well as segments with electronic noise were excluded from analyses.

Once textgrids were generated for each task and condition, a Praat script was used to automate data extraction from the audio recordings. The functions of the script relevant to the present study are as follows:

1) Pairs of sound files and textgrids with matching file names were read into Praat.
2) The script settings were specified to analyze a) 5 formants within the spectrogram; b) disregard all information above 5,500 Hz within the signal (i.e., the maximum frequency range), and c) use a pitch floor of 75 Hz and ceiling of 1000 Hz in

calculations to accommodate possible unanticipated and extreme changes to frequency as a result of increased vocal SPL in the loud speaking conditions.

3) The script ran through each textgrid looking for *words* on one tier (i.e., the breath group tier); these were individually extracted as *parts* (smaller audio clips) using rectangular windows given the start and end points of each *word* (breath group) segment.
4) From the smaller parts, the script next looked for *segments* boundaries on another tier (i.e., the phon tier). These were also extracted as parts for analysis (again, using rectangular windows).
5) F0 values were extracted for each segment. We used mean F0 to guide analyses of dB SPL, where vowels with pitch tracking errors were discarded from analyses.
6) Intensity values were extracted using 0.005s windows and a pitch floor of 100 Hz. The analyses for the current study used mean intensity values extracted from 100% of the segment duration. A pitch floor of 100 Hz was chosen to improve F0 tracking and avoid error outputs as a result of low frequency segments, such as those with creaky or breathy quality.
7) Values are printed to a text file along with the time point, vowel label, and breath group label at which the value was extracted.

Mean dB SPL values were calculated from vowels across each breath group for each participant in each experimental task and condition. Once all means were obtained, the 440 Hz calibration tone was read into Praat and a dB SPL value was obtained for a 1 second sample taken as close to the sound level meter reading as possible. The difference between Praat-generated measurements and the output value recorded from the sound level meter during microphone calibration (actual SPL) was calculated. This difference value was subsequently added to all Praat-generated dB SPL measures in order to accurately reflect the vocal SPLs produced during experimental tasks (Boersma & Weenink, 2011).

Chest wall kinematic and lung volume analysis

Kinematic and integrated flow signals from were digitized (1000 Hz) and low pass filtered (3 Hz), and displayed in LabVIEW v7.0 (National Instruments) along with the digitized acoustic record. To guide analyses, the video record of the study session was simultaneously displayed on a television monitor. Data were analyzed following the protocol outlined in Boliek, Hixon, Watson, & Jones (2009), in which signals were configured in a *y-t* (motion-time) and *y-x* format (motion-motion). The *y-x* format was used to establish end expiratory level (EEL), which is the reference point for zero volume displacement. In this study, EEL was defined as the minimum value for one tidal breathing cycle before each breathing task. Kinematic signals were subsequently referenced to this EEL, such that volumes above EEL were expressed as positive values and those below EEL were expressed as negative values.

The custom software program described in Boliek, Hixon, Watson, & Jones (2009) was used to analyze the kinematic data in the present study. Video recordings were reviewed several times to ensure that only stable breath groups were selected for analyses. Calibrated rib cage and abdominal volume initiations, terminations, and excursions were used to derive the following measurements for each breath group during speech and nonspeech breathing tasks: 1) Lung volume at breathing task initiation in mL (LVI); 2) Lung volume at breathing task termination in mL (LVT); 3) Lung volume excursion used during the breathing task, which was calculated as the difference between LVI and LVT and expressed as a percent of predicted vital capacity

(%LVE); and 4) Percent rib cage contribution to lung volume excursion (%RC) during the breathing task, which was calculated by dividing rib cage excursion by LVE.

Surface electromyographic and coherence analysis

The EMG waveforms from the intercostal (IC) and oblique (OB) muscles were selected for analysis from the expiratory limb of the breath group (from onset to offset of expiration) as determined by the summed and calibrated kinematic signals representing lung volume (see Figure 2). The digitized acoustic record was presented on a separate channel and was used to verify that the data selection included speech produced by the participant. These data were then passed through a Tukey window to reduce erroneous high-frequency signals at the borders of adjoining breath trials, rectified and concatenated, for subsequent coherence analysis. The average total window length of the concatenated data used for coherence analysis was 75.95 for phonation, 29.25 for sentence repetition, and 53.02 for storytelling.

Intermuscular coherence analysis was performed for the two chest wall muscle sites (EMG_{IC}-EMG_{OB}). We report the amplitude of coherence and the frequency where significant coherence occurs. Intermuscular coherence was calculated as:

$$\boldsymbol{MSC} = \left|\boldsymbol{C_{xy}(w)}\right|^{2} = \frac{\left|\overline{\boldsymbol{G_{xy}(w)}}\right|^{2}}{\overline{\boldsymbol{G_{xx}(w)}} \cdot \overline{\boldsymbol{G_{yy}(w)}}}$$

Where MSC denotes magnitude square coherence; Gxx(w) and Gyy(w) are the averaged power spectra of the x and y muscles of interest (here, the IC and OB muscles) for a given frequency (w). Gxy represents the averaged cross-power spectrum of x and y signals at frequency w (from Halliday et al., 1995). Segments of length 2048 points with no overlap were derived from a frequency resolution of 2.44 Hz and calculated offline in the MATLAB environment

(MathWorks, Natick, MA) using open access scripts (`www.neurospec.org`) (Halliday et al., 1995). Data were analyzed for coherence within two frequency bandwidths: Low (15-59 Hz) and high (60-110 Hz). These ranges were chosen to remove potential overlap between data points and accounted for the 2.44 Hz frequency resolution used in MATLAB. Significant coherence in each of these bands was defined as the peak correlation identified from a minimum of 3 sequential data points (+/- 5 Hz) above the 95% confidence interval (see Figure 3). The 95% confidence interval for significant non-zero coherence was calculated as .0332 for phonation, .0521 for sentence repetition, and .02818 for storytelling.

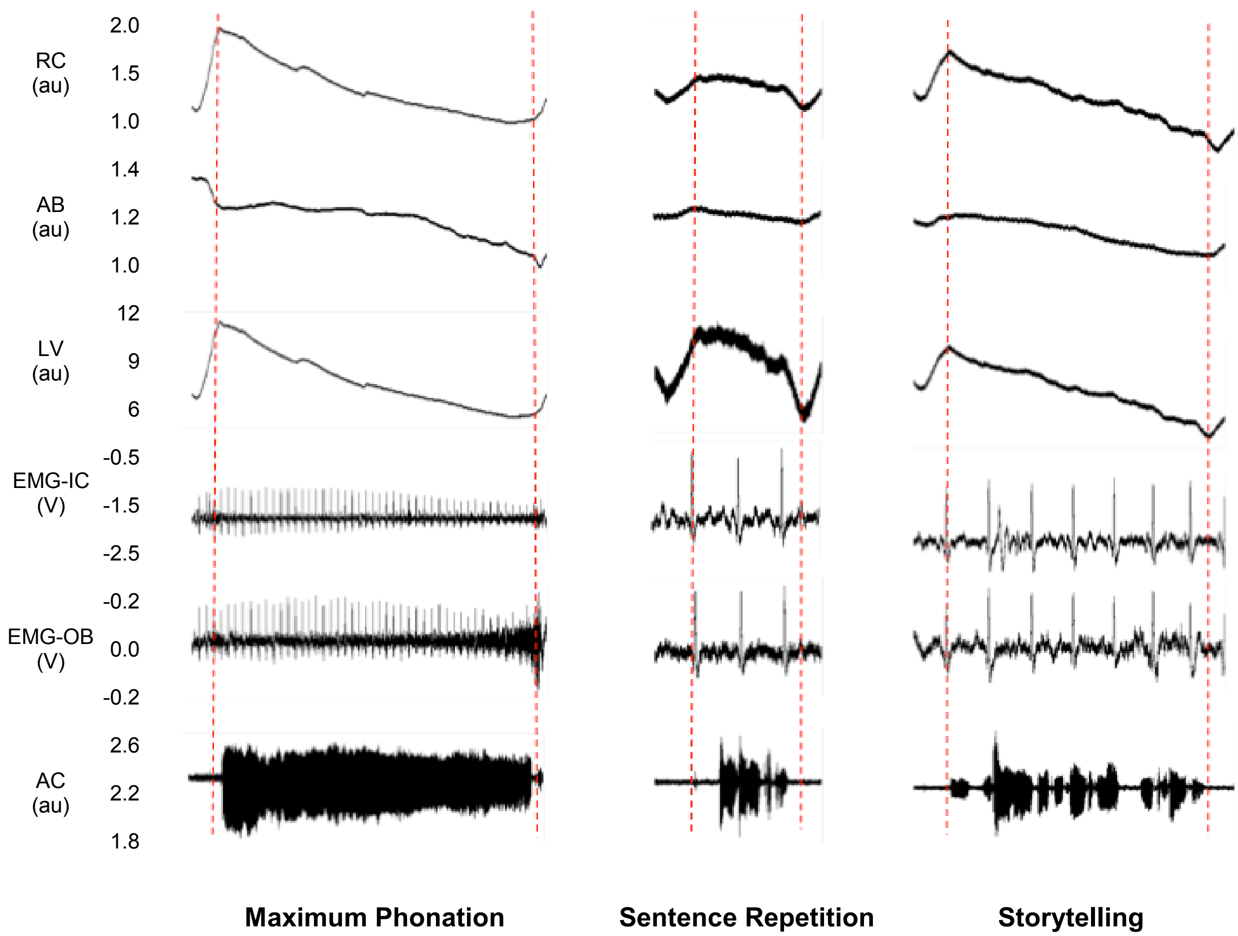


Fig 2. Raw waveforms data showing rib cage (RC) and abdominal (AB) kinematic signals, lung volume (LV) corrected for respective RC and AB contributions (see Methods section), intercostal (IC) and oblique (OB) EMG activity, and the acoustic signal (AC) during all experimental tasks for a representative participant. Dotted lines denote the onset (A) and offset (B) of expiration, as measured by the LV signal. au = arbitrary units, V = volts.

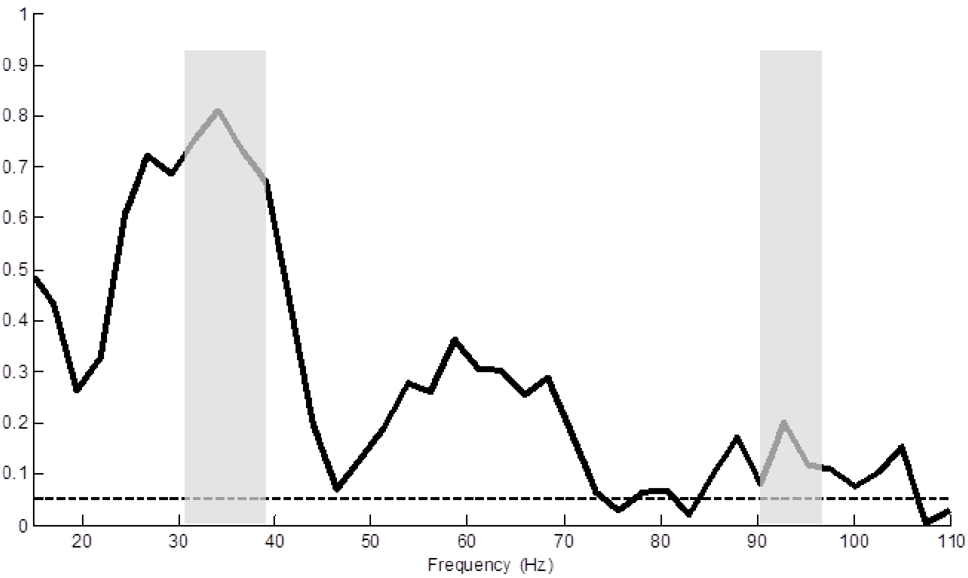


Fig 3. Coherence spectra tracings during sentence repetition for a representative participant. The 95% confidence limit at .05496 is shown. Shaded regions represent the low frequency (15-59 Hz) and high frequency (60-110 Hz) areas of peak coherence.

Interrater Reliability

For each dependent variable, ten percent of the data were randomly selected and analyzed to assess interrater reliability. Interclass correlations and their 95% confidence intervals were calculated using a single-rating, absolute-agreement, two-way mixed effects model (SPSS Statistical package version 23, SPSS Inc., Chicago, IL). Reliability was considered excellent (greater than .90), good (between .75 and 9), moderate (.5 and .75), and poor (less than .5) based on the 95% confidence interval of the ICC value (Koo & Li, 2016). Table 2 reports the interclass correlation coefficients (ICC) for each measurement of interest and task.

Measurement	Task	Interrater reliability	ICC	95% CI
dB SPL	Phonation	Excellent	.991	[.928, 1]
	Sentence repetition	Excellent	.995	[.928, 1]
	Storytelling	Excellent	.999	[.986, 1]
LVI	Phonation	Good	.878	[-.099, .992]
	Sentence repetition	Good	.886	[-.199, .992]
	Storytelling	Moderate	.590	[-1.06, .97]
LVT	Phonation	Excellent	.987	[.866, .999]
	Sentence repetition	Good	.836	[-.296, .989]
	Storytelling	Excellent	.960	[-.033, .998]
%RC	Phonation	Excellent	.995	[.927, 1]
	Sentence repetition	Moderate	.572	[-2.97, .971]
	Storytelling	Moderate	.691	[-3.01, .980]
%LVE	Phonation	Excellent	.994	[.918, 1]
	Sentence repetition	Good	.783	[-.399, 985]
	Storytelling	Good	.803	[-.851, .987]
Coherence (Low frequency band)	Phonation	Excellent	.999	[.988, 1]
	Sentence repetition	Excellent	.996	[.960, 1]
	Storytelling	Excellent	.998	[.978, 1]
Coherence (High frequency band)	Phonation	Excellent	.953	[.655, .995]
	Sentence repetition	Good	.795	[-1.23, .979]
	Storytelling	Moderate	.613	[-5.75, .962]

Table 2. Interclass correlation coefficients (ICC) and the 95% confidence interval (95% CI) of the ICC values as a measurement of interrater reliability for 10% of the data.

Intrarater Reliability

For measurements made by each rater, ten percent of the data were randomly selected and reanalyzed by the same rater to assess intrarater reliability.

For measurements of dB SPL, excellent reliability was found within raters. Average ICC for measurements made by rater one was .995, 95% CI [.928, 1]. Average ICC for measurements made by rater two was .999, 95% CI [.986, 1].

For measurements of lung volume, excellent reliability was found within raters. Average ICC for measurements made by rater one was .997, 95% CI [.993, .999]. Average ICC for measurements made by rater two was .996, 95% CI [.987, .999].

For measurements of chest wall intermuscular coherence, excellent reliability was found within raters. Average ICC for measurements made by rater one was .972, 95% CI [.757, .996]. Average ICC for measurements made by rater two was .997, 95% CI [.993, .999].

Results

For each task (maximum phonation, sentence repetition, storytelling), separate one-way repeated measures analyses of variance (ANOVA) were used to examine the effects of cue type (CONV, CUED, NOISE) on variables related to the modulation of vocal loudness (vocal SPL, intermuscular coherence, lung volume events). Bonferroni corrections were applied to main effects and post hoc analyses. When the assumption of homogeneity of variance was violated, Greenhouse-Geisser corrected values were reported.

Effect of cue on vocal SPL

A total of 8754 acoustic segments (vowels) met inclusion criteria and were submitted to statistical analyses (135 maximum phonations, 2904 vowels for sentence repetition, 5715 vowels for storytelling). The F ratios and p values for the repeated measures ANOVAs are shown in Table 3. Table 4 summarizes the mean and standard deviation effects of cue on vocal SPL, as well as results from post-hoc paired t-tests performed when significant main effects were found. Across all tasks, greater vocal SPLs were produced in the CUED and NOISE conditions than the CONV condition. No significant differences in mean vocal SPL were found between the CUED and NOISE conditions during maximum phonation, sentence repetition, or storytelling, indicating that instruction to speak at "twice conversational loudness" and presentation of multi-talker noise elicited similar changes to vocal SPL. Figure 4 depicts results from post-hoc paired samples t-tests indicating mean differences in vocal SPL between each condition across all tasks. Speakers were exposed to an estimated noise level between 70 to 91.28 dB SPL (M = 82.72, SD = 6.51) in the NOISE condition.

Task	*F* ratio	*p* value	*df*
Phonation	28.95	< .001*[a]	1.37, 19.13
Sentence repetition	57.66	< .001*	2,28
Storytelling	56.08	< .001*	2,28

Table 3. *F* ratios and probability values for repeated measures ANOVAs on the effect of cue on mean vocal SPL. Asterisks denote significant effects at $p \leq .017$. [a]Greenhouse-Geisser value reported.

	CONV		CUED		NOISE				
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	CUED-CONV	NOISE-CONV	NOISE-CUED
Phonation	82.00	4.07	90.38	4.77	87.03	4.83	$t_{14} = 12.39$, $p < .001$*	$t_{14} = 4.35$, $p < .001$*	$t_{14} = 2.44$, $p = .029$
Sentence repetition	79.23	2.22	86.58	3.16	84.82	3.07	$t_{14} = 10.06$, $p < .001$*	$t_{14} = 8.87$, $p < .001$*	$t_{14} = 2.27$, $p = .39$
Storytelling	81.79	2.68	86.79	2.96	87.36	2.41	$t_{14} = 9.15$, $p < .001$*	$t_{14} = 9.48$, $p < .001$*	$t_{14} = .94$, $p = .365$

Table 4. Results from post-hoc paired samples *t*-tests (two-tailed) and probability values comparing the effect of cue on mean vocal SPL for each task. Asterisks denote significant differences at $p \leq .017$.

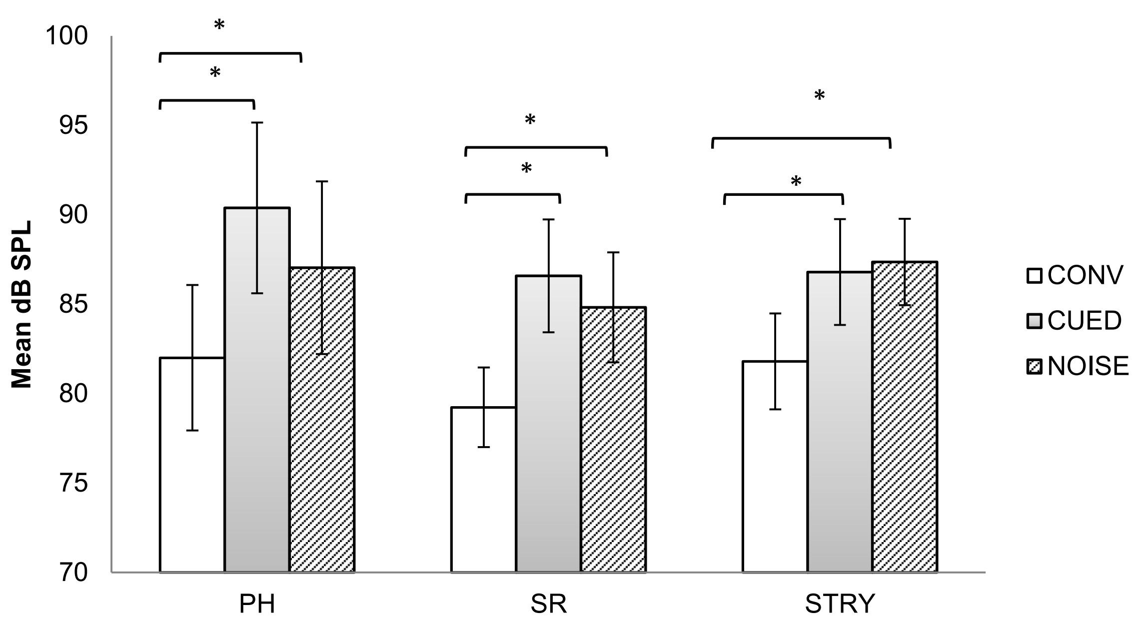


Fig 4. Results from post-hoc paired samples *t*-tests (two-tailed) indicating significant differences in mean vocal SPL between cue conditions for each task. Error bars represent standard deviation. Asterisks denote significant differences at p ≤ .017. PH = maximum phonation, SR = sentence repetition, STRY = storytelling.

Effect of cue on lung volume

A total of 1077 breath groups met inclusion criteria and were subject to statistical analyses (135 maximum phonations, 403 breath groups for sentence repetition, and 539 breath groups for storytelling). Table 5 reports F ratios and p values for the repeated measures ANOVAs. Results revealed a significant main effect of condition on %LVE during sentence repetition only. No other significant main effects of condition were found ($p > .004$). Post hoc paired samples t-tests with a corrected alpha level of $p \leq .017$ indicated that %LVE used during sentence repetition was significantly greater in the CUED condition than in the CONV condition ($t = 4.81$, $df = 14$, $p < .001$, two-tailed). Greater %LVE was also used in the NOISE condition than in the CONV condition ($t = 4.02$, $df = 14$, $p = .001$, two-tailed). Finally, %LVE was significantly different between loud conditions during sentence repetition, with greater %LVE used in the CUED condition than in NOISE ($t = 3.21$, $df = 14$, $p = .006$, two-tailed). Table 6 summarizes the mean and standard deviation effects of cue on speech breathing measurements. Figure 5 depicts the results of one-way repeated measures ANOVAs on lung volume variables derived from chest wall kinematics, as well as results from post hoc tests indicating differences in %LVE between cue conditions during sentence repetition as described above.

Measurement	Task	*F* ratio	*p* value	*df*
LVI	Phonation	.08	.925	2,28
	Sentence repetition	4.15	.026	2,28
	Storytelling	.69	.508	2,28
LVT	Phonation	.78	.467	2, 28
	Sentence repetition	.16	.851	2,28
	Storytelling	5.12	.013	2,28
%LVE	Phonation	.97	.393	2,28
	Sentence repetition	18.95	< .001*[a]	1.4, 19.65
	Storytelling	2.65	.115	1.29,18
%RC	Phonation	.69	.510	2,28
	Sentence repetition	2.51	.121[a]	1.39,19.46
	Storytelling	2.29	.120	2,28

Table 5. *F* ratios and probability values for repeated measures ANOVAs on lung volume at utterance initiation (LVI), lung volume at utterance termination (LVT), lung volume excursion used during the breathing task expressed in percent vital capacity (%LVE), and percent rib cage contribution to lung volume change (%RC). Asterisks denote significant effects at $p \leq .004$. [a]Greenhouse-Geisser value reported.

			CONV		CUED		NOISE	
Measurement	**Unit**	**Task**	***M***	***SD***	***M***	***SD***	***M***	***SD***
LVI	mL	Phonation	2241.13	707.08	2183.20	792.60	2196.90	919.93
		Sentence	545.76	288.14	834.04	533.18	700.52	296.15
		Storytelling	878.81	446.82	843.65	353.46	757.33	368.73
LVT	mL	Phonation	-1282.83	993.02	-1284.13	1112.55	-1435.89	1125.65
		Sentence	12.87	228.40	58.92	426.77	21.97	257.70
		Storytelling	175.99	380.99	-1.11	411.74	-47.92	355.48
%LVE	%VC	Phonation	75.21	29.41	73.78	32.43	77.45	27.57
		Sentence	11.27	4.31	16.73	4.65	14.74	4.28
		Storytelling	14.90	5.16	17.88	5.88	17.51	8.76
%RC	%RC	Phonation	80.36	7.55	78.77	8.49	79.14	9.3
		Sentence	76.43	9.18	80.20	11.61	75.34	15.01
		Storytelling	81.47	10.21	81.84	12.54	78.14	10.94

Table 6. Mean and standard deviation effects of cue on lung volume by cue condition and task. Measurements include lung volume at utterance initiation (LVI), lung volume at utterance termination (LVT), lung volume excursion used during the breathing task expressed in percent vital capacity (%LVE), and percent rib cage contribution to lung volume change (%RC). Negative values in LVT indicate volumes below end expiratory level.

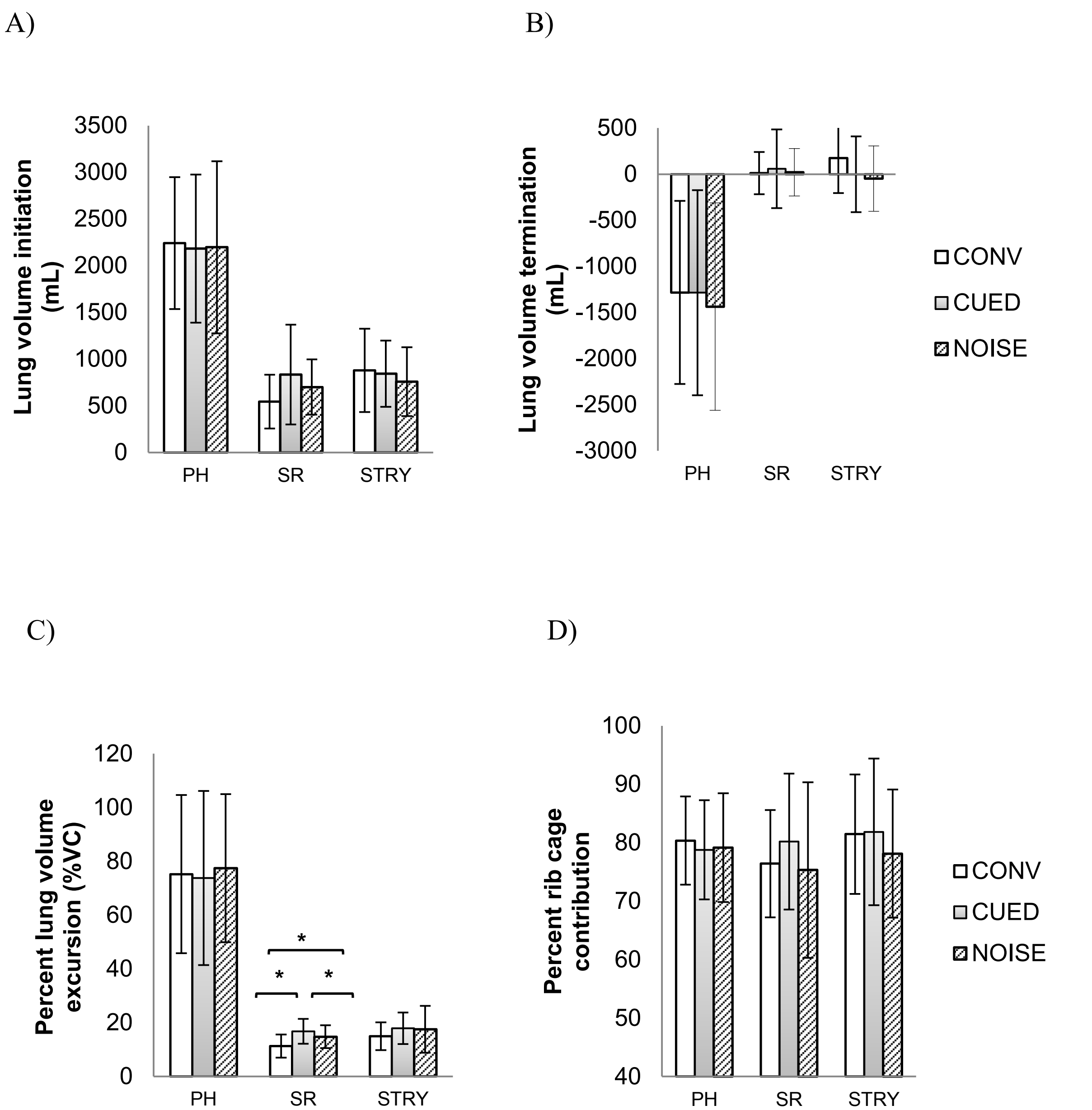


Fig. 5. Results of one-way repeated measures ANOVAs testing for effects of cue on lung volume variables derived from chest wall kinematics: A) Lung volume initiation (mL); B) Lung volume termination (mL); C) Percent lung volume excursion expressed as percent vital capacity (%VC); D) Percent rib cage contribution to lung volume excursion. Error bars represent standard deviation. Asterisks denote significance at ($p \leq .017$). PH = phonation, SR = sentence repetition, STRY = storytelling.

Effect of cue on chest wall intermuscular coherence

A total of 1077 breath groups analyzed met inclusion criteria and were subject to statistical analyses (135 maximum phonations, 403 breath groups for sentence repetition, and 539 breath groups for storytelling). Prior to group-level analyses, peak coherence measures were converted to Fisher's z-scores in order to allow for comparisons between correlation values. Table 7 reports F ratios and p values for the repeated measures ANOVAs. No significant main effects of condition were found for peak coherence in either the high or low frequency bandwidths ($p > .008$). Figure 6 depicts the results from the repeated measures ANOVAs used to test for effects of cue on transformed coherence values in the low and high frequency bandwidths. Table 8 summarizes the mean and standard deviation effects of cue on peak coherence values, expressed in terms of correlation (i.e., untransformed values). Table 9 reports the frequency at which mean peak coherence occurred in each frequency band for each task and condition.

Paired samples t-tests were used to test for differences between transformed mean peak coherence values between the low and high frequency bandwidths for each task (see Figure 7). Results showed that, across all tasks, peak amplitudes of coherence were significantly greater in the low frequency band than in the high frequency band during phonation for $p \leq .017$ (CONV: $t_{14} = 6.80$, $p < .001$; CUED: $t_{14} = 6.57$, $p < .001$, NOISE: $t_{14} = 6.44$, $p < .001$); sentence repetition (CONV: $t_{14} = 9.77$, $p < .001$, CUED: $t_{14} = 10.21$, $p < .001$, NOISE: $t_{14} = 11.21$, $p < .001$); and storytelling (CONV: $t_{14} = 8.54$, $p < .001$, CUED: $t_{14} = 8.11$, $p < .001$, NOISE: $t_{14} = 8.75$, $p < .001$). Figure 10 in Appendix A depicts individual responsiveness to cue condition across each task.

Frequency band	Task	*F* ratio	*p* value	*df*
Low	Phonation	.14	.867	2,28
	Sentence repetition	2.45	.104	2,28
	Storytelling	2.81	.077	2,28
High	Phonation	2.42	.108	2,28
	Sentence repetition	1.77	.20[a]	1.41, 19.68
	Storytelling	1.60	.221	2,28

Table 7. *F* ratios and probability values for repeated measures ANOVAs on significant (above the 95% confidence limit) peak coherence values in two frequency bands: Low (15-59 Hz) and High (60-110 Hz). [a]Greenhouse-Geisser values reported.

		CONV		CUED		NOISE	
Frequency Band	**Task**	***M***	***SD***	***M***	***SD***	***M***	***SD***
Low	Phonation	.49	.24	.51	.24	.50	.24
	Sentence repetition	.71	.14	.65	.16	.71	.11
	Storytelling	.62	.18	.56	.19	.58	.21
High	Phonation	.15	.14	.11	.10	.10	.11
	Sentence repetition	.28	.21	.22	.17	.22	.15
	Storytelling	.17	.14	.17	.14	.13	.15

Table 8. Mean and standard deviation effects of cue on raw peak coherence values in the low (15-59 Hz) and high (60-110 Hz) bands.

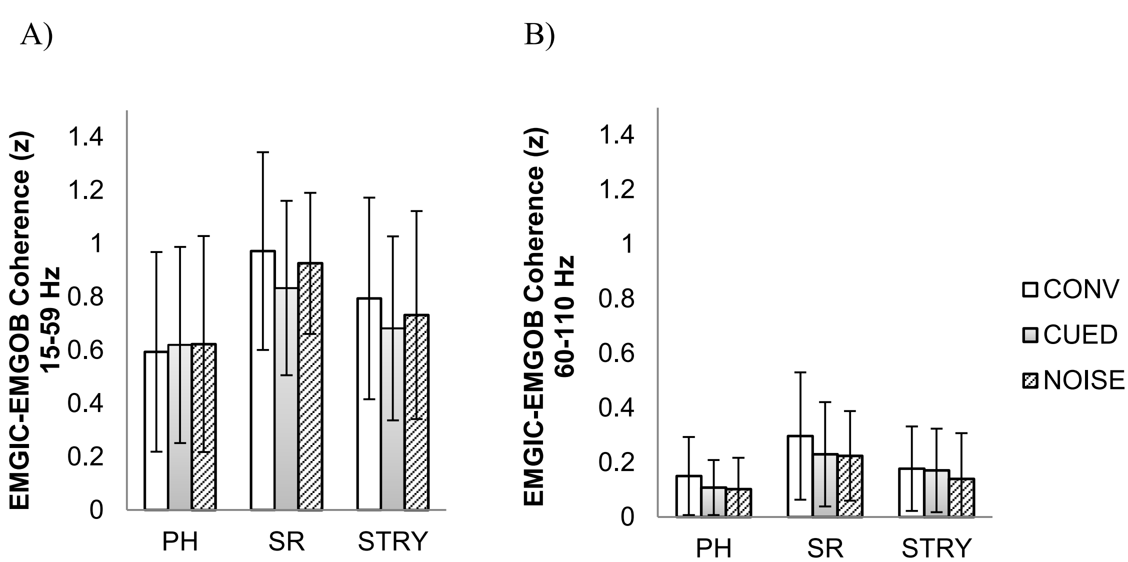


Fig. 6. Results from one-way repeated measures ANOVAs on transformed mean peak coherence values (*z*) for each task in the low (A, 15-59 Hz) and high (B, 60-110 Hz) frequency bands. No significant main effects found ($p > .008$). Error bars represent standard deviation.

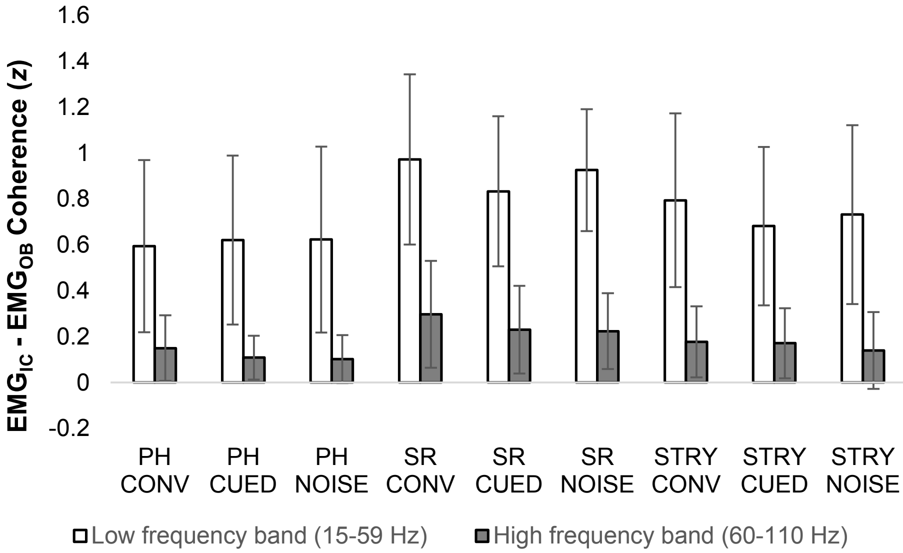


Fig 7. Results from paired samples *t*-tests (two-tailed) comparing transformed mean peak coherence values (*z*-scores) between the low and high frequency bandwidths for each task and cue condition. All comparisons are significant at $p \leq .017$. Error bars represent standard deviation. PH = maximum phonation; SR = sentence repetition; STRY = storytelling.

Frequency band	Task	CONV (Hz)	CUED (Hz)	NOISE (Hz)
Low	Phonation	76.81	78.13	68.36
	Sentence repetition	71.47	76.81	73.85
	Storytelling	73.24	77.75	73.91
High	Phonation	30.27	30.27	30.52
	Sentence repetition	31.90	30.27	31.74
	Storytelling	32.55	31.25	33.53

Table 9. Summary of the frequencies (Hz) at which mean peak coherence occurred in the low (15-59 Hz) and high frequency bandwidths (60-110 Hz) for each task and condition across all subjects who demonstrated nonzero coherence (significant at the 95% confidence interval).

Exploratory Analyses

Laryngeal activity. Because there are several ways in which the speech mechanism can produce vocal loudness changes, we recognized the possibility that lack of significant effects was due to simultaneous recruitment of, or reliance on, laryngeal mechanisms for vocal loudness adjustment. To test the possibility that adjustments to vocal SPL were made with the laryngeal system, we further examined the effects of vocal loudness cue on fundamental frequency (F_0). F_0

was calculated from the same vowels and using the same Praat script as in the dB SPL analysis. Descriptive analyses using the Shapiro-Wilk test in conjunction with measures of standardized skewness indicated that the F_0 data met the assumption of normality. Therefore, we analyzed the effect of cue on F_0 using a one-way repeated measures ANOVA without separating our sample by gender to reduce the risk of Type I error and to increase statistical power. Table 10 reports F ratios and p values for the repeated measures ANOVAs. Figure 8 summarizes the mean and standard deviation effects of cue on F_0, as well as results from post-hoc paired t-tests (two-tailed) performed when significant main effects were found. During phonation, mean F_0 was significantly higher in the CUED ($t = 5.34$, $df = 14$, $p < .001$) and NOISE ($t = 3.02$, $df = 14$, $p = .009$) conditions than in the CONV condition. No differences in mean F_0 were found between the loud speaking conditions ($p = .22$). A similar pattern was observed during storytelling, in which mean F_0 values in the CUED ($t = 5.88$, $df = 14$, $p = < .001$) and NOISE ($t = 4.84$, $df = 14$, $p = < .001$) conditions were significantly greater than in the CONV condition. No differences were observed in mean F0 between the loud speaking conditions ($p = .61$) However, a different pattern in mean F_0 change was observed during sentence repetition; as with the other tasks, mean F_0 was significantly higher in the CUED condition ($t = 2.99$, df =14, $p = .01$) as compared with the CONV condition; however, mean F_0 produced in the NOISE condition and CONV condition did not differ ($p = .086$). In addition, the sentence repetition task produced differences in F_0 between the loud speaking conditions, with significantly higher mean F_0 in the CUED condition ($t = 3.42$, $df = 14$, $p = .004$). This difference can be attributed to the similarity in variability of F_0 between the CUED and NOISE conditions, in which relatively small changes in F_0 constituted a significant difference.

Chest Wall Muscle Activity. Previous work has used measurements of rib cage and abdominal volume displacement to assert that speakers use increased muscular effort when instructed to speak "twice as loud" relative to increasing vocal SPL under the Lombard effect (Huber, Chandrasekaran, & Wolstencroft, 2005). However, volume displacements do not necessarily reflect muscle activity because lower volumes can also be achieved by passive recoil forces. Because EMG offers a more direct measurement of muscular effort, we conducted an exploratory analysis to investigate potential effects of cue on IC and OB muscle activity. One-way repeated measures ANOVAs (Table 10) showed no significant differences in chest wall muscle activity across conditions. Table 11 summarizes the mean and standard deviation effects of cue on IC and OB muscle activity, expressed as a percentage of maximum voluntary contraction.

Task	Measurement	*F* ratio	*p* value	*df*
Phonation	F_0	10.63	< .001*	2,28
	%MVCIC	.11	.725	2,28
	%MVCOB	.05	.953	2,28
Sentence repetition	F_0	6.74	.015*[a]	1.22, 17.02
	%MVCIC	2.05	.167 [a]	1.31,18.28
	%MVCOB	3.34	.050	2,28
Storytelling	F_0	20.79	< .001*	2,28
	%MVCIC	2.98	.067	2,28
	%MVCOB	.44	.651	2,28

Table 10. *F* ratios and probability values for repeated measures ANOVAs on mean fundamental frequency (F_0), intercostal (%MVCIC), and oblique muscle activity (%MVCOB), expressed as a percentage of maximum voluntary contraction. Asterisks denote significant effects at $p \leq .017$.
[a]Greenhouse-Geisser value reported.

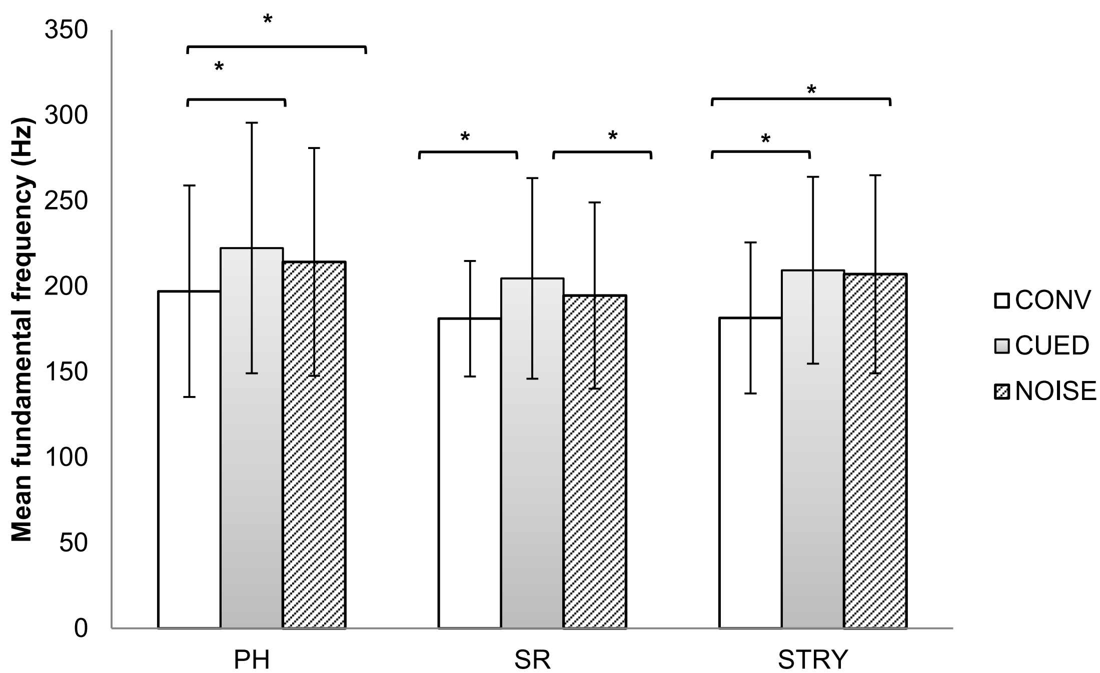


Fig 8. Results from post-hoc paired samples *t*-tests comparing mean fundamental frequency produced in each task and cue condition. Asterisks denote significance at $p \leq .017$. Error bars represent standard deviation. PH = maximum phonation; SR = sentence repetition; STRY = storytelling.

		CONV		CUED		NOISE	
Measurement	**Task**	***M***	***SD***	***M***	***SD***	***M***	***SD***
%MVCIC	Phonation	32.42	31.97	33.34	29.52	36.07	37.07
	Sentence repetition	6.34	12.85	12.68	9.90	11.17	15.19
	Storytelling	12.90	13.35	22.23	18.50	12.31	13.85
%MVCOB	Phonation	34.28	27.80	34.15	32.87	35.58	31.47
	Sentence repetition	4.38	5.41	7.58	8.10	4.65	6.92
	Storytelling	11.51	14.12	14.92	14.48	12.41	15.33

Table 11. Mean and standard deviation effects of cue on intercostal and oblique muscle activity, expressed as a percentage of maximum voluntary contraction (%MVCIC, %MVCOB, respectively) for each task.

Relationship between vocal SPL and physiological mechanisms. To better understand the relationships between vocal SPL and the physiological mechanisms involved in vocal loudness adjustment, we conducted a bivariate correlation analysis between dB SPL and four parameters that represent the biomechanics involved in vocal loudness adjustment for each condition and task. The biomechanical parameters included indirect measures of laryngeal adjustment (F_0); the lung volumes that speakers inspired to in order to support speech at greater SPLs (LVI), and the expiratory muscular effort involved in generating pressure for loud speech (%MVCIC, %MVCOB). Table 12 shows the resulting correlations coefficients (Pearson's *r*). There were no significant relationships between dB SPL and any of the variables during phonation. During sentence repetition, a positive relationship was observed between dB SPL and %MVCOB during in the CUED condition ($p = .034$), as well as between dB SPL and LVI in the NOISE condition ($p = .009$). During storytelling, there was a positive correlation between dB SPL and %MVCOB in the CUED condition ($p = .027$).

Task	Cue Condition	F_0	LVI	%MVCIC	%MVCOB
Phonation	CONV	-.115	-.042	.059	.455
	CUED	-.21	.025	.247	.366
	NOISE	.187	.054	.169	-.079
Sentence Repetition	CONV	.46	-.208	-.19	.192
	CUED	-.151	.369	.106	.55*
	NOISE	.493	.645**	-.481	.078
Storytelling	CONV	.091	-.002	.196	.136
	CUED	-.249	.16	.264	.568*
	NOISE	.26	.407	.196	.442

Table 12. Bivariate correlations (Pearson's *r*) between mean dB SPL and four main outcome variables representing biomechanics of vocal loudness adjustment: Fundamental frequency (F_0), lung volume at utterance initiation (LVI), intercostal and oblique muscle activity expressed as a percentage of maximum voluntary contraction (%MVCIC, %MVCOB, respectively). Correlations that reached significance are denoted by an asterisk. $^*p < .05$. $^{**}p < .01$

Intermuscular coherence during speech and nonspeech breathing tasks. Previous work has shown that coherence in the high frequency 60-110 Hz range is suppressed during speech production relative to controlled deep breathing (Smith & Denny, 1990; Denny & Smith, 2000), whereas coherence in the low 20-60 Hz range remains comparable across speech and nonspeech breathing tasks (Smith & Denny, 1990). We conducted a comparison of peak coherence values obtained during breathing at rest and the storytelling task in order to examine whether or not we could replicate these findings. The storytelling task in the CONV condition was chosen for this analysis because it was most comparable to the reading task used in Smith and Denny (1990). In line with these previous findings, paired samples *t*-tests revealed that mean peak coherence in the high frequency band was significantly greater during nonspeech resting breathing (M = .45, SD = .21) than during the storytelling task (M = .18, SD = .15) t = 3.85, df = 14, p = .002, two-tailed. However, differences also were observed in the low frequency band between tasks, with greater mean peak coherence occurring during resting breathing (M = 1.1, SD = .45) than during storytelling (M = .79, SD = .38) t = 2.77, df = 14, p = .015, two-tailed (Figure 9).

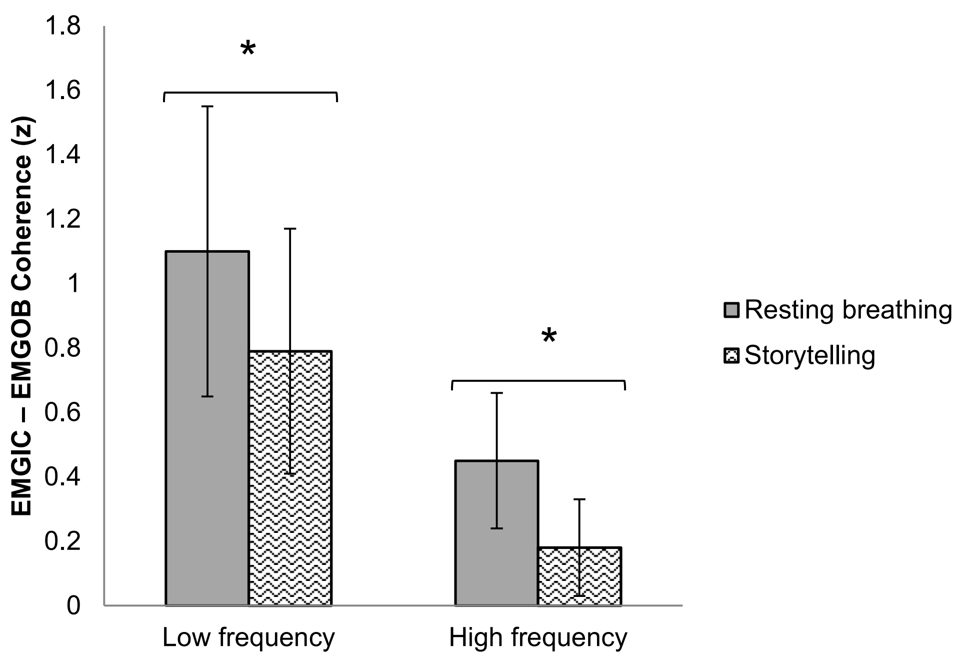


Fig 9. Results from post hoc paired samples *t*-tests comparing transformed mean peak coherence values (*z*) produced during resting breathing (non-speech task) and storytelling in the CONV condition (speech task) in the low frequency band (15-59 Hz) and the high frequency band (60-110 Hz). Asterisks denote significance at $p \leq .025$. Error bars represent standard deviation.

Discussion

The primary purpose of this study was to examine whether patterns of speech breathing control, as measured by acoustics, chest wall kinematics, and intermuscular coherence would vary according to different cues for vocal loudness. The main findings were: 1) Similar increases to vocal SPL were produced in response to cues for vocal loudness; 2) Given similar increases to vocal SPL, different cues did not result in systematic changes to speech breathing parameters; 3) For each task, low frequency coherence (15-59 Hz) was comparable across all cues, and high frequency coherence (60-110 Hz) was similar across cues, but greater mean peak coherence was observed in the low frequency bandwidth than in the high frequency band, regardless of cue or task. Exploratory analyses indicated that speakers likely modulated vocal loudness using laryngeal adjustments, as measured by increased F_0 in the loud speaking conditions.

Effect of cue on vocal SPL

Across all tasks, mean vocal SPL increased by approximately 5 to 8 dB above regular conversational loudness in the CUED and NOISE conditions. This magnitude of vocal SPL change is similar to mean values reported in other studies that have modified speakers' vocal SPL using instructions to speak "twice as loud" as typical conversation (Dromey & Ramig, 1998; Huber, Chandrasekaran, & Wolstencroft, 2005; Huber, 2007), as well as those that have elicited the Lombard effect using similar levels of noise (e.g., Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Garnier, Henrich, & Dubois, 2010). Compared with other studies however, we aimed to elicit a +7 dB SPL increase in order to control for potential variability as a result of differences in tracheal pressure, thus facilitating our ability to make comparisons between our physiological measurements across conditions. Because speakers produced similar changes to vocal SPL in our loud conditions, the results from this study and others suggest that healthy adults are fairly consistent in their responses when instructed to produce phonation or speech at twice conversational loudness.

Regarding the Lombard effect, presentation of multi-talker noise resulted in significant increases to vocal SPL relative to conversation, thereby eliciting changes to vocal loudness that are characteristic of the phenomenon. The level of noise presentation was specific to each speaker and ranged between 70 to 91.28 dB SPL, which is comparable to the range of noise presentation levels chosen by other studies that have investigated the Lombard effect (e.g., Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Huber, Chandrasekaran, & Wolstencroft, 2005; Huber, 2007; Patel & Schell, 2008; Garnier, Henrich, & Dubois, 2010). The mean noise presentation level in the current study of approximately 83 dB SPL is particularly similar to the level chosen by Cooke and Lu (2010), who presented 82 dB SPL of multi-talker noise through headphones in order to ensure adequate energetic as well as informational masking effects (Cooke & Lu, 2010).

Effect of cue on lung volume

One aim of this study was to replicate previous work reporting that speech breathing behaviours systematically change as a function of vocal loudness cue. In the current study, speakers did not adopt consistent cue-specific breathing patterns across tasks, despite significant increases to vocal SPL. Effects of cue were restricted to percent lung volume excursion (%LVE) during sentence repetition, with the greatest mean proportion of lung volume (16.73 %VC) used in the CUED condition. To produce sentences at “twice conversational loudness”, speakers expended approximately 5% more of their available lung volume as compared to the CONV condition. The NOISE condition also resulted in a significant, but relatively smaller change to lung volume excursion of over 3% VC (i.e., the Lombard effect) during sentence repetition. In comparison, previous work has reported that the Lombard effect produces the greatest %LVE during sentence reading in noise, with a mean change in lung volume of approximately 2 %VC from conversational loudness (Huber, Chandrasekaran, & Wolstencroft, 2005).

Apart from these findings, no other significant differences were observed between cueing conditions on selected measures of speech breathing. During phonation and storytelling, lung volumes at utterance initiation and termination, as well as proportions of lung volume and rib cage movement were not affected by the different cues. Our null findings during the storytelling task conflicts with Huber (2007) and Huber (2008), who reported significantly greater lung volume initiations, terminations, and excursions during monologue production in noise than at conversational SPL. Huber (2007) provided a comparison of results across related investigations that have explored modifications to lung volume in response to different loudness cues. Table 13 updates these findings with results from the current study and shows relevant measurements and comparisons from similar populations and tasks.

The absence of a relationship between lung volume and cueing condition may be related to several methodological factors. First, the present study employed a novel approach to examining respiratory responses to different loudness cues, in which we deliberately held the loudness target constant between the CUED and NOISE conditions (approximately 7 dB SPL above conversational loudness). Other studies comparing the Lombard effect to instructed vocal loudness change have not employed such a control and have reported that speakers use significantly greater vocal SPL during the Lombard effect than when instructed to speak at twice conversational loudness (Huber, 2007; Stathopoulos et al., 2014). As such, previous reports of cue-based respiratory strategies during dynamic speech production are difficult to interpret, in that differences in tracheal pressure targets alone could have driven changes to speech breathing patterns. Importantly, by ensuring that speakers produced equivalent increases to vocal SPL, we may have resolved any potential effects of cue on speech breathing parameters during spontaneous speech.

Measurement	**CUED**	**NOISE**
Lung volume initiation	--[A, B, D]	--[A, B, E] ↑[C, F]
Lung volume termination	--[A, B, D]	--[A, B, C, E] ↑[F]
Lung volume excursion	↑[A*] --[B, D]	↑[A*, B, C, F] --[E]
Percent rib cage contribution	--[A] ↑[B]	--[A] ↑[B]

Table 13. Comparison of findings across various studies investigating the effect of different cues on selected parameters of speech breathing. -- denotes no change from baseline conversational loudness level; ↑ denotes increase. All studies reviewed here used multi-talker noise to elicit the Lombard effect. Only the current study and Winkworth & Davis (1997) delivered noise through headphones; all other studies delivered noise free field through speakers. With the exception of the current study, all studies delivered noise at a single level to all participants.

A. Current study. Asterisks denote that findings were significant for sentence repetition only. No effects were observed during storytelling.
B. Huber, Chandrasekaran, & Wolstencroft, 2005: CUED and NOISE during sentence repetition
C. Huber, 2007: NOISE during monologue production.
D. Dromey & Ramig, 1998: CUED during word repetition. Did not report on percent rib cage contribution.
E. Winkworth & Davis, 1997: NOISE during monologue production. Did not report on percent rib cage contribution.
F. Huber, 2008: NOISE during monologue production. Did not report on percent rib cage contribution.

Another methodological consideration concerns the effects of utterance length on lung volume. Recall that lung volume is related to the length of an utterance: Longer utterances have been shown to promote greater lung volumes at utterance initiation, lower lung volumes at utterance termination, and greater proportions of lung volume used during loud and conversational-level spontaneous speech (Winkworth, Davis, Adams, & Ellis, 1995; Huber, 2008). Further, lung volumes at utterance initiation have been shown to be related to those at the termination of the previous utterance (Winkworth, Davis, Adams, & Ellis, 1995). Based on these findings, it is not surprising that we found a statistical trend in the lung volumes used during the sentence repetition task (%LVE), in which utterances were limited and consistent in length. That is, it is possible that the storytelling task introduced variability in lung volumes as a result of variability in utterance length, thereby making it difficult to identify statistical significance in these data. Indeed, our data indicate that storytelling in the NOISE condition produced the most variability in %LVE (as measured by standard deviation) as compared with the CUED and CONV conditions. Our data also show that the %LVE used during storytelling was slightly more variable than during sentence repetition, across all conditions. Further support for this explanation comes from inconsistent findings across studies that investigate the effect of cue on monologue production in healthy adults. Whereas some studies have reported significant effects of noise on %LVE only (Huber, 2007; Sadagopan & Huber, 2007), other have reported that background multi-talker babble significantly increased measures of LVI, LVT, and %LVE from those observed at baseline conversational loudness (Huber, 2008). Again however, vocal SPL was not controlled for in these other studies.

Investigations of chest wall kinematics also have shown that speakers use significantly lower abdominal initiation and termination volumes during “twice as loud” speech than in response to other cues (Huber, Chandrasekaran, & Wolstencroft, 2005), which has been used to substantiate the claim that intentionally increasing vocal SPL requires greater muscular effort and is less biomechanically efficient than increasing vocal SPL under the Lombard effect. Interpreting abdominal volume displacement as an indicator of active muscular effort is problematic however, because displacement can occur as a result of passive recoil. In the current study, we had access to direct measurements of muscle activation through use of surface EMG. Our results showed that different vocal loudness cues did not influence the magnitude of intercostal or oblique muscle EMG measurements, indicating that respiratory muscle activity remained stable during vocal loudness adjustments. During exploratory analyses however, we found that expiratory muscular activity in the abdominal wall increased with vocal SPL during sentence repetition and storytelling in the CUED condition, thus demonstrating a relationship between abdominal muscle activity and intentional changes to vocal loudness. As discussed, contraction of the abdominal wall is integral to biomechanical efficiency of the respiratory system during vocal loudness modulation (Hixon, Goldman, & Mead, 1976; Stathopoulos & Sapienza, 1997).

Effect of cue on chest wall intermuscular coherence

The primary aim of this study was to investigate chest wall intermuscular coherence produced during vocal loudness change. Previous work has identified that correlated oscillations in two frequency bands are common to respiratory activity: 20-60 Hz and 60-110 Hz (Smith & Denny, 1990). These bands are thought to represent drive to muscles of the speech mechanism from distinct areas of the central nervous system: Coherent activity in the lower frequency band has been associated with common drive from the corticospinal tract, likely originating from the motor cortex (Grosse, Cassidy, & Brown, 2002; Hansen et al., 2005; Norton & Gorassini, 2006; Fisher, Zaaimi, Williams, Baker, & Baker, 2013), whereas coherence in the high frequency band is thought to represent rhythmic output from medullary central pattern generators (CPGs) subserving respiration (Bruce & Ackerson, 1986; Smith & Denny, 1990; Denny & Smith, 2000). Based on this understanding of speech breathing control and the assumption that the Lombard effect is primarily mediated by brainstem feedback mechanisms (Nonaka, Takahashi, Enomoto, Katada, & Unno, 1997; Hage, Jürgens, & Ehret, 2006), we predicted that the Lombard effect would promote coherence peaks in the high frequency bandwidth. However, we found that coherence between the intercostal and oblique muscles remained stable across all frequency bandwidths studied, regardless of cue type. Since peaks were observed in both the high and low frequency bandwidths, our results seem to suggest that chest wall intermuscular coherence produced during vocal loudness adjustment is consistent across comparable vocal SPL measurements in healthy adult speakers, irrespective of the cue used to elicit increased vocal SPL.

We also predicted a significant reduction in peak amplitude of chest wall coherence during the CUED condition on the basis of previous work, which has shown that β-band coherence diminishes during tasks that require divided attention (Stepp, Hillman, & Heaton, 2010) and increased precision (Kristeva-Feige, Fritsch, Timmer, & Lücking, 2002). Instruction to speak "twice as loud" in the CUED condition was presumed to require greater attention and precision control over vocal SPL than the relatively natural cue of multi-talker noise. Our results did not support this hypothesis, as there was no significant effect of cue on the mean peak amplitude of chest wall coherence in either of the frequency bands studied.

Interpreting these coherence data in the context of our chest wall kinematic and muscle activation findings can offer additional insights. In this study, physiological measurements of the chest wall were largely unperturbed by different cues: Patterns of chest wall intermuscular coherence, as well as the lung volumes and muscular activations used during vocal loudness adjustment remained constant as speakers increased vocal loudness in response to instruction as well as in response to multi-talker noise. That is, a main finding of this work was that when the level of vocal loudness output was controlled for, the biomechanical behaviours of the respiratory system during vocal loudness change and the underlying respiratory neuromuscular control mechanisms used to drive them were similarly unaffected by different cues to increase loudness. Our results therefore offer converging evidence for the stability of respiratory control networks used to drive vocal loudness adjustment in healthy young adults.

Although not included in our original hypotheses, we found greater peak coherence values in the low frequency band than the high frequency band for each task and cue condition (see Figure 5). In addition, our coherence values in the low frequency band appear to be higher than those previously reported in the literature for similar tasks. In the current study, average raw peak coherence values in the low frequency band ranged from r = .65 to .71 during sentence repetition and from r = .55 to .62 during storytelling, whereas Smith and Denny (1990) reported that maximum 20-110 Hz chest wall coherence ranged from .03 and .25 for eight subjects during paragraph reading at a single "moderately" loud level. Several methodological differences between the current study and Smith and Denny (1990) may account for the dissimilarity between the coherence values reported between the two studies. The first difference relates to the protocol for ground electrode placement. In our study, the site of the ground was the clavicle, whereas Smith and Denny (1990) recorded reference activity from the right first dorsal interosseus muscle as a baseline condition. Current recommendations on use of surface EMG from muscles of speech and swallowing specify that a ground electrode should be placed close to the recording sites of interest over electrically neutral tissue, such as a bony prominence (Stepp, 2012, also see De Luca, 2002). It is likely that use of a control condition as opposed to a ground electrode placed at an appropriate site in Smith and Denny (1990) resulted in relatively poorer signal-to-noise ratio and potentially limited ability to detect muscular activity as compared to the current study. Second, we recorded respiratory activity from the 6^{th} intercostal space and the oblique muscle, whereas Smith and Denny (1990) obtained bilateral EMG recordings from the 7^{th} and 8^{th} intercostal spaces. Recordings from different combinations of chest wall muscles may have led to different coherence values, especially if patterns of respiratory muscle activity differed between studies. A related point is that the current study analyzed the expiratory limb of speech tasks for significant peak intermuscular coherence between expiratory muscles, because speech is produced on expiration. From their protocol, it is unclear whether Smith and Denny (1990) analyzed the inspiratory limb, expiratory limb, or both in their speech task (passage reading). Because their aim was to detect coherence from the diaphragm, which is the primary muscle of inspiration (Hixon, Weismer, & Holt, 2014), then the selection of EMG signals should only have been made from the inspiratory limb of the speech task. Finally, it may be the case that intermuscular coherence values differ between spontaneous speech production, which was elicited in the current study, and reading aloud, which was the speech task used in Smith and Denny (1990). It has been well-established that breathing behaviours used during oral reading are different from those used during spontaneous speech (e.g., Hodge & Rochet, 1989; Winkworth, Davis, Ellis, & Adams, 1994; Winkworth, Davis, Ellis, & Adams, 1995; Mitchell, Hoit, & Watson, 1996; Huber, 2007; Wang, Green, Nip, Kent, & Kent, 2010). Given that intermuscular coherence also has been shown to be sensitive to speech task (Stepp, Hillman, & Heaton, 2010), it seems reasonable to suggest that differences in coherence values observed between the current study and Smith and Denny (1990) were potentially a consequence of the tasks used to elicit speech production.

The primary aim of the current study was to compare effects of different cues to increase loudness on chest wall coherence within a series of speech breathing tasks. In contrast, the focus of previous work on coherence between respiratory muscles has been directed towards differentiating the control mechanisms involved in speech versus nonspeech breathing tasks (Smith & Denny, 1990; Denny & Smith, 2000). In their work, Smith and Denny (1990) and

Denny and Smith (2000) reported a pattern in which low frequency coherence was comparable across speech (reading aloud) and nonspeech breathing tasks (tracking a visually-presented wave to elicit controlled deep breathing), whereas high frequency coherence (60-110 Hz) was absent or significantly diminished during speech relative to non-speech breathing tasks. This pattern of coherence was used to assert that a necessary tradeoff in distinct respiratory control mechanisms occurs during speech production, in which the brainstem circuits used to drive metabolic breathing are suppressed to accommodate the respiratory requirements for speech. Although not included in our primary aim, we conducted an exploratory analysis to examine whether or not we could replicate this pattern of findings. Consistent with previous data, our data showed significantly greater peak coherence in the high frequency band during resting breathing relative to storytelling at conversational loudness. However, we did not reproduce the finding that speech and nonspeech breathing tasks produce similar low frequency coherence, because greater peak coherence was observed during resting breathing relative to storytelling in the low frequency band (see Figure 9).

It is possible that we found task-related differences in low frequency peak coherence values because we used a spontaneous speech task, whereas Smith and Denny (1990) and Denny and Smith (2000) used a reading task. As discussed above, breathing behaviours during spontaneous speech tasks have been shown to be different than those used during oral reading; these differences have been interpreted to reflect higher cognitive-linguistic demands of spontaneous speech (Mitchell, Hoit, & Watson, 1996; Wang, Green, Nip, Kent, & Kent, 2010). Given that factors related to task difficulty and demands on cognitive skill have been shown to reduce the strength of peak coherence in the low frequency band (Stepp, Hillman, & Heaton, 2010), it may be the case that the cognitive-linguistic load associated with spontaneous language formulation drove the pattern seen in our coherence results that have not been previously observed. Compared with oral reading, spontaneous speech tasks better reflect the demands of everyday communication; therefore, our results could be considered to be a more valid representation of the neuromuscular control patterns used during speech production.

Although the number of studies investigating high frequency intermuscular coherence in respiratory muscles during speech is limited, it is difficult to reconcile our results within a dichotomous nonspeech-speech framework of understanding high frequency coherence. Because significant (i.e., non-zero) high frequency coherence has been shown to occur during speech breathing tasks in this study and others (Smith & Denny, Denny & Smith, 2000), it is not likely that its appearance strictly reflects central pattern generator control over nonspeech breathing. This notion is further supported Denny and Smith (2000), who reported that nearly half of their typically fluent speakers demonstrated equivalent low (20-60 Hz) and high frequency coherence (60-110 Hz) between chest wall muscles during their speech task, suggesting that suppression of high frequency coherence is not necessary during speech.

Contributions of laryngeal mechanisms to vocal loudness adjustment

Studies that have made concurrent measurements of laryngeal and respiratory functions have shown that for some speakers, laryngeal mechanisms predominate the control of vocal SPL (Stathopoulos & Sapienza, 1993). In the current study, information about laryngeal involvement was limited to measurement of fundamental frequency (F_0), which is considered an indirect acoustic measure of intrinsic laryngeal muscle activation (Ohala, 1970; Titze, 1988). Specifically, elongation of the cricothyroid muscle has been identified as the primary means of increasing F_0 (Ohala, 1970; Hirano, 1974). Our analysis of mean F_0 suggested at least some involvement of laryngeal musculature during vocal loudness adjustment: For the phonation and storytelling tasks, significantly higher mean F_0 was observed in the CUED and NOISE conditions than in the CONV condition. During sentence repetition, instructions to speak "twice as loud" not only resulted in significantly higher mean F_0 than in the CONV condition, but in the NOISE condition as well. These findings are in line with Dromey & Ramig (1998) who reported that F_0 increased significantly when they instructed speakers to repeat a word at "twice" and "four times" their typical conversational loudness. Similarly, increased F_0 has been considered a hallmark acoustic feature of the Lombard effect (Patel & Schell, 2008; Cooke & Lu, 2010; Garnier, Henrich, & Dubois, 2010).

Limitations

In the current study, insert earphones were used as part of a noise delivery system designed to ensure that speakers produced comparable increases to vocal SPL in the loud speaking conditions. Although this system was successful in eliciting similar changes to vocal SPL, it is possible that delivering noise through earphones or headphones introduced variability into speech breathing patterns, resulting in a lack of significant findings in our physiological variables. Similar to this study, Winkworth & Davis (1997) delivered multi-talker noise through supra-aural headphones and also found that speakers did not use consistent speech breathing strategies during the Lombard effect. In contrast, studies that have reported cue-dependent kinematic patterns have either used loud speakers to deliver free field multi-talker noise (Huber, Chandrasekaran, & Wolstencroft, 2005; Huber, 2007) or used a modified in-the-ear hearing aid to deliver multi-talker noise monaurally (Stathopoulos et al., 2014). It appears that the effect of noise presentation method is not confined to the respiratory system because articulatory movements also have been shown to vary according to whether noise is delivered through headphones or loudspeakers (Garnier, Henrich, & Dubois, 2010).

One possible explanation for these findings could be that free field or open ear noise delivery methods better represent the auditory feedback conditions that speakers encounter in natural Lombard conditions. Use of insert earphones likely modified speakers' perceptions of auditory feedback as a result of the *occlusion effect*, which refers to the enhanced intensity of bone conducted and low frequency sounds that occurs when the ear canal is occluded (Hood, 1962). The communicative environment itself could also have been considered unnatural, since speakers wore headphones and the experimenter did not. If speakers in the current study perceived the auditory and/or communicative environment as unnatural or altered as a result of wearing earphones, then they could have deliberately employed a variety of compensatory strategies to overcome the effects of the earphones, leading to the absence of consistent, cue-specific patterns in respiratory behaviour and its underlying control. Moreover, our data support the possibility that speakers found speaking in both conditions equally as difficult because similar values of low frequency coherence in the CUED and NOISE conditions could reflect comparable modulation of coherence between chest wall muscles as a function of task difficulty. To these points, comments from two speakers in this study offer important anecdotal evidence: One speaker commented that he purposely spoke quieter in the NOISE condition because he was unable to hear himself. Another reported feeling embarrassed during the NOISE condition and consciously attempted to lower the loudness of her voice. For at least two speakers then, factors related to the naturalness of the Lombard effect influenced behaviour in the NOISE condition. These anecdotal reports are in line with findings from one study in which speakers indicated greater discomfort and poorer perception of their own voice when speaking in noise delivered through headphones than when speaking in free-field noise delivered over loud speakers (Garnier, Henrich, & Dubois, 2010). It is interesting that our physiological data suggest that speakers engaged in voluntary override of the Lombard effect, despite our acoustic data indicating that speakers were unable to do so.

With regards to calibration of our noise delivery system, the coupler-derived measurements of dB SPL did not account for individual variability in ear canal dimensions; therefore, the exact level of sound output delivered to individual speakers in this study could not be determined. A more precise evaluation of sound output level could have been achieved with real ear measurements, which depict the actual level of acoustic energy within an individual's ear canal (Humes & Bess, 2014). However, obtaining real ear measurements required that a registered audiologist to be present for each experimental session, and such measurements were beyond the scope of the current work.

Finally, focusing on respiratory parameters may have not adequately described all of the physiological adaptations underlying the Lombard effect and may have contributed to our lack of significant findings on measures of lung volume, chest wall muscle activity, and chest wall intermuscular coherence. In our exploratory analyses, we found that measurements of F_0 increased in the loud speaking conditions. Although these data suggested that speakers recruited laryngeal mechanisms to change vocal SPL, it has been shown that vocal fold tension can increase passively as a result of displacement from midline with increasing SPL (Titze, 1988). Therefore, it is unclear as to what extent speakers actively recruited laryngeal mechanisms during vocal loudness adjustment. In addition to laryngeal mechanisms, speakers could also have manipulated oral articulatory parameters to facilitate effective transmission of sound energy to the atmosphere, thereby reducing impedance and increasing vocal SPL (Hixon, Weismer, & Holt, 2014). Previous work has shown that upper and lower lip displacements increase with vocal SPL when speakers are asked to speak at "twice" and "four times" their conversational loudness (Dromey & Ramig, 1998). Similar observations have been made for the Lombard effect, in which speakers demonstrate higher values of the first formant frequency and greater mouth opening when speech is produced in noise than at conversational loudness in quiet conditions (Garnier, Henrich, & Dubois, 2010).

Future Directions

Future studies aiming to use the Lombard effect as a naturalistic manipulation of vocal loudness may need to consider the effects of noise delivery system on measurements of interest. To better simulate natural speech in noise conditions, one solution could be to play additional feedback of a speaker's voice into their headphones in order to compensate for potential attenuation or effects of occlusion (Garnier, Henrich, & Dubois, 2010; Cooke & Lu, 2010). Another solution could be positioning the listener further away from the listener to improve the naturalness of the communication environment. The listener could also wear headphones to simulate comparable listening conditions to the speaker. Additionally, future studies could deliver noise through loudspeakers and apply noise cancellation processing techniques (Ternstrom, Bohman, & Sodersten, 2006), although use of these techniques may be limited by technical expertise in this area. Studies that have used loud speakers to investigate of the effect of noise on respiratory kinematics have not outlined the methods with which noise was removed from the acoustic signal (Huber, Chandrasekaran, & Wolstencroft, 2005; Huber, 2007); however, separating noise from the acoustic signal may not have been a concern for these studies that were not primarily interested in acoustic measurements. Finally, use of noise delivery method as an independent variable for manipulating respiratory, laryngeal, and oral articulatory variables during speech production warrants further research.

Analyses of laryngeal contributions to vocal loudness adjustment also could help to interpret the results in the current study, which focused on the respiratory system in isolation. Because F_0 is considered an indirect measure of intrinsic laryngeal muscular activity, future research aiming to examine laryngeal responses to loudness cues could use aerodynamic measures known to directly reflect laryngeal activity during vocal loudness change (e.g., open quotient of the glottis, maximum flow declination rate of the glottal waveform). Conducting coherence analyses of the cricothyroid muscles could potentially offer a more direct window into neuromuscular control of the larynx; however, it should be noted that cricothyroid activity would not likely be detected from surface EMG recordings because of its relatively deep position in the neck. Instead, signals would likely reflect strap muscle activation (Stepp, 2012). In addition, further analyses of formant structure and spectrum features from our acoustic data could provide insight on changes to articulatory parameters in response to vocal loudness cues. To our knowledge, there have not been any studies that have taken simultaneous measurements of oral-articulatory, laryngeal, and respiratory functioning during the Lombard effect in healthy young adults, although some recent work has been done in this area in healthy older adults, as well as in those with Parkinson's disease (Matheron, Stathopoulos, Huber, & Sussman, 2017; Stathopoulos et al., 2014).

Much of the recent work on speech motor control and the Lombard effect has focused on establishing whether respiratory support of vocal loudness is more efficient in noise, particularly in individuals who demonstrate low vocal SPL secondary to Parkinson's disease (Sadagopan & Huber, 2007; Stathopoulos et al., 2014; Huber, Stathopoulos, & Sussman, 2014). Our data suggest that biomechanical efficiency is not a consequence of the Lombard effect at least in healthy young adults whose speech systems likely represent optimal conditions of neuromuscular control. Therefore, the present work provides comparative data for understanding changes to speech motor control as a result of aging, as well as processes related to the pathophysiology of neuromuscular disorders, including Parkinson's disease. A complementary line of questioning then, is to examine whether and how patterns of lung volumes and intermuscular coherence vary with different cues in children, who may be less efficient at manipulating the respiratory system for speech than adults. During the school-age years, the systems and strategies used to regulate speech breathing undergo a period of refinement (Boliek, Hixon, Watson, & Jones 2009). Compared with adults, younger children (under 10 yrs) make changes to vocal SPL primarily with the breathing apparatus, do not take advantage of passive recoil forces at higher lung volumes, and initiate speech using even greater inward movement of the abdomen (Stathopoulos and Sapienza, 1997). One proposal for age-related differences in pressure, volume, and shape variables used for speech is maturation of the neural control and coordination of the chest wall. Others contend that greater airway compliance in childhood enables speech production at higher or lower lung volume levels without the increased effort and muscular work required in an adult system (Russell & Stathopoulos, 1988; Solomon & Charron, 1998). Though the Lombard effect has been reported to occur in children as young as 3 and 4 years of age (Siegel, Pick, Olsen, & Sawin, 1976), there are no known reports on chest wall kinematics during naturalistic manipulation of vocal loudness among children. It is possible that the Lombard effect represents a less effortful and more efficient adjustment to vocal loudness in the developing speech and language system. Therefore, replicating this work in young children and interpreting results in the context of the adult data from the present study may offer valuable insight into development of neuromotor control over speech breathing. There also may be clinical implications. The Lombard effect has recently been shown to be a more biomechanically efficient adjustment of vocal loudness in adults with dysphonia secondary to Parkinson's disease (Stathopoulos et al., 2014); as a result, it has been applied as a treatment to promote healthy vocal loudness during everyday conversation (Huber, Stathopoulos, & Sussman, 2014). Insofar as Lombard adjustments for healthy vocal loudness are biomechanically efficient in children, then this future work could provide a platform for investigating application of the effect to children with motor speech disorders.

Conclusions

To our knowledge, this was the first study to investigate modulation of chest wall intermuscular coherence using different vocal loudness cues. The main finding of this study was that different cues had no effect on the control of speech breathing, indicating that the respiratory control circuits involved in vocal loudness change are distributed and stable against perturbation in the healthy adult system when the targeted sound pressure level remains similar between different cues to increase loudness. The results add to a growing body of literature in which the application of surface EMG and coherence analyses are used as a noninvasive means for understanding speech motor control.

References

Amazi, D.K., & Garber, S.R., (1982). The Lombard sign as a function of age and task. *Journal of Speech and Hearing Research, 25,* 581-585.

Alberta College of Speech-Language Pathologists and Audiologists. (2015). *Hearing screening guideline: Preschool to adult.* Retrieved from https://acslpa.ab.ca/download/college/Hearing%20Screening%20Guideline.pdf

American Speech-Language-Hearing Association Panel on Audiologic Assessment. (1997). *Guidelines for Audiologic Screening*. Retrieved from http://www.asha.org/policy/GL1997-00199/.

Boliek, C.A., Hixon, T.J., Watson, P.J., & Jones, P.B. (2009). Refinement of speech breathing in healthy 4- to 6-year-old children. *Journal of Speech, Language & Hearing Research, 52*(4), 990-1007.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer, version 6.0.19 [Computer software]. Retrieved from www.praat.org.

Boersma, P., & Weenink, D. (2011). Sound pressure calibration. *Praat: Doing Phonetics by Computer.* Retrieved from http://www.fon.hum.uva.nl/praat/manual/sound_pressure_calibration.html.

Boonstra, T.J. (2013). The potential of corticomuscular and intermuscular coherence for research on human motor control. *Frontiers in Human Neuroscience, 7,* 1-2.

Bruce, E.N., & Ackerson, L.M. (1986). High frequency oscillations in human electromyogram s during voluntary contractions. *Journal of Neurophysiology, 56,* 542-553.

Brumm, H., & Zollinger, S.A. (2011). The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour, 148,* 1173-1198.

Clair-Auger, J.M., Gan, L.S., Norton, J.A., & Boliek, C.A. (2015). Simultaneous measurement of breathing kinematics and surface electromyography of chest wall muscles during maximum performance and speech tasks in children: Methodological considerations. *Folia Phoniatrica et Logopaedica*, *67*, 202-211.

Cooke, M., & Lu, Y. (2010). Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. *Journal of the Acoustical Society of America, 128*(4), 2059-2069.

Denny, M., & Smith, A., (2000). Respiratory control in stuttering speakers: Evidence from respiratory high frequency oscillations. *Journal of Speech, Language, and Hearing Research, 43,* 1024-1037.

DeLuca, C.J. (2002). Surface electromyography: Detection and Recording. *DelSys Incorporated.* 1-10.

Dromey, C., & Ramig, L.O. (1998). The effect of lung volume on selected phonatory and articulatory variables. Journal of Speech, Language & Hearing Research, 41(3), 491-502.

Duffy, J.R. (2013). *Motor Speech Disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Lousi, Missouri: Elsevier Mosby.

Eliades, S.J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology, 89*, 2194-2207.

Eliades, S.J., & Wang, X. (2005). Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex, 15,* 1510-1523.

Eliades, S.J., & Wang, X. (2012). Neural correlates of the Lombard effect in the primate auditory cortex. *The Journal of Neuroscience, 32*(21), 10737-10748.

Fisher, K.M., Zaaimi, B., Williams, T.L., Baker, S.N., Baker, M.R. (2012). Beta-band intermuscular coherence: a novel biomarker of upper motor neuron dysfunction in motor neuron disease. *Brain, 135,* 2849-2864.

Garnier, M., Henrich, N., & Dubois, D. (2010). Influence of sound immersion and communicative interaction on the Lombard effect. *Journal of Speech, Language, and Hearing Research, 53*(3), 588-608.

Grosse, P., Cassidy, M.J., Brown, P. (2002). EEG-EMG, MEG-EMG and EMG-EMG frequency analysis: physiological principles and clinical applications. *Clinical Neurophysiology*, *113*, 1523-1531.

Hage, S.R., Jürgens, U., & Ehret, G. (2006). Audio-vocal interaction in the pontine brainstem during self-initiated vocalization in the squirrel monkey. *European Journal of Neuroscience, 23,* 3297-3308.

Halliday, D.M., Rosenberg, J.R., Amjad, A.M., Breeze, P., Conway, B.A., & Farmer, S.F. (1995). A framework for the analysis of mixed time series/point process data-theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. *Progress in Biophysics and Molecular Biology, 64*(2-3), 237-278.

Hansen, N.L., Conway, B.A., Halliday, D.M., Hansen, S., Pyndt, H.S., Biering-Sørensen, F., & Nielsen, J.B. (2005). Reduction of common synaptic drive to ankle dorsiflexor motoneurons during walking in patients with spinal cord lesion. *Journal of Neurophysiology, 94,* 934-942.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatrica et logopaedica, 26,* 89-94.

Hixon, T., Goldman, M., & Mead, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of Speech and Hearing Research*, *16*, 78-115.

Hixon, T.J., Mead, J., & Goldman, M.D. (1976). Dynamics of the chest wall during speech production: function of the thorax, rib cage, diaphragm, and abdomen. *Journal of Speech, Language, and Hearing Research, 19*(2), 297-356.

Hixon, T.J., & Weismer, G. (1995). Perspectives on the Edinburgh study of speech breathing. *Journal of Speech, Language & Hearing Research, 38*(1), 42-60.

Hixon, T.J., Weismer, G., & Holt, J.D. (2014). *Preclinical speech science* (2nd ed.). San Diego, CA: Plural Publishing Inc.

Hodge, M.M., & Rochet, A.P. (1989). Characteristics of speech breathing in young women. *Journal of Speech, Language, and Hearing Research, 32*(3), 466-480.

Hoit, J.D., & Hixon, T.J., (1987). Age and speech breathing. *Journal of Speech, Language, and Hearing Research, 30*(3), 351-66.

Hoit, J.D., Hixon, T.J., Watson, P.J., & Morgan, W.J. (1990). Speech breathing in children and adolescents. . *Journal of Speech, Language & Hearing Research, 33*(1), 51-69.

Hood, J.D. (1962). Bone conduction: A review of the present position with especial reference to Dr. Georg von Békésy. *The Journal of the Acoustical Society of America, 34*(8), 1325-1332.

Houde, J.F., Nagarajan, S.S, Sekihara, K., & Merzenich, M.M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14*(8), 1125-1138.

Huber, J.E. (2007). Effect of cues to increase sound pressure level on respiratory kinematic patterns during connected speech. *Journal of Speech, Language, and Hearing Research, 50,* 621-634.

Huber, J.E., Chandrasekaran, B., & Wolstencroft, J.J. (2005). Changes to respiratory mechanisms during speech as a result of different cues to increase loudness. *Journal of Applied Physiology, 98*(6), 2177-2184.

Huber, J. Stathopoulos, E. & Sussman J. (2014). Changes to respiratory and laryngeal function after SpeechVive treatment in individuals with Parkinson's disease. Paper presented at the 2014 Convention of the American Speech-language and Hearing Association, Orlando, FLA.

Humes, L.E., & Bess, F.H. (2014). *Audiology and communication disorders: An overview* (2nd Ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Isshiki, N. (1964). Regulatory mechanism of vocal intensity variation. *Journal of Speech, Language, and Hearing Research, 7,* 17-29.

Jaiser, S.R., Baker, M.R., Baker, S.N. (2016). Intermuscular coherence in normal adults: Variability and changes with age. *PLoS ONE, 11*(2), 1-16.

Jurafsky, D., Bell, A., Fosler-Lussier, E., Girand, C., Raymond, W. (1998). Reduction of English function words in Switchboard. In R. H. Mannell & J. Robert-Ribes, (Eds.). *Proceedings from ICSLP '98: 5th International Conference on Spoken Language Processing Proceedings, 7,* p. 3111. Canberra City, Australia: Australasian Speech Science and Technology Association.

Kanai, T., & Wang, S.C. (1962). Localization of the central vocalization mechanism in the brain stem of the cat. *Experimental Neurology, 6*(5), 426-434.

Kilner, J.M., Baker, S.N., Salenius, S., Jousmaki, V., Hari, R., & Lemon, R.N. (1999). Task-dependent modulation of 15-30 Hz coherence between rectified EMGs from human hand and forearm muscles. *Journal of Physiology, 15,* 559-570.

Kilner, J.M., Baker, S.N., Salenius, S., Hari, R., & Lemon, R.N. (2000). Human cortical muscular coherence is directly related to specific motor parameters. *Journal of Neuroscience, 20*(23), 8838-8845.

Konno, K., & Mead, J. (1967). Measurement of the separate volume changes of the rib cage and abdomen during breathing. *Journal of Applied Physiology, 22,* 407-422.

Kristeva-Feige, Fritsch, C., Timmer, J., & Lücking, C.H. (2002). Effects of attention and precision of exerted force on beta range EEG-EMG synchronization during a maintained motor contraction task. *Clinical Neurophysiology, 113,* 124-131.

Ladefoged, P., (1963). Some physiological parameters in speech. *Language and Speech, 6*(3), 109-119.

Ladefoged, P., & McKinney, N., (1963). Loudness, sound pressure and sub-glottal pressure in speech. *Journal of the Acoustical Society of America, 35,* 454-460.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech, Language, and Hearing Research, 14,* 677-709.

Lehmann, G.J., & McGill, S.M. (1999). The importance of normalization in the interpretation of surface electromyography: a proof of principle. *Journal of Manipulative and Physiological Therapeutics, 22*(7), 444-446.

Lu, Y., & Cooke, M. (2009). The contribution of changes in F0 and spectral tilt to increased intelligibility of speech production in noise. *Speech Communication, 51*(12), 1253-1262.

Matheron, D., Stathopoulos, E.T., Huber, J.E., Sussman, J, E. (2017). Laryngeal aerodynamics in healthy older adults and adults with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 60*(3), 494-507.

Maurer, C., von Tscharner, V., Nigg, B.M. (2013). Speed-dependent variation in the Piper rhythm. *Journal of Electromyography and Kinesiology, 23,* 673-678.

Milenkovic, P. (2001). TF32 [computer software]. Madison, WI: University of Wisconsin–Madison.

Mitchell, H.L., Hoit, J.D., & Watson, P.J. (1996). Cognitive-linguistic demands and speech breathing. *Journal of Speech, Language, and Hearing Research, 39*(1), 93-104.

Netsell, R., & Hixon, T.J. (1978). A noninvasive method for clinically estimating subglottal air pressure. *Journal of Speech and Hearing Disorders, 43, 326-330.*

Koo, T.K., & Li, M.Y. (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

Nonaka, S., Takahashi, R., Enomoto, K., Katada, A., & Unno, T. (1997). Lombard reflex during PAG-induced vocalization in decerebrate cats. *Neuroscience Research, 29*(4), 283-289.

Norton, J.A., & Gorassini, M.A. (2006). Changes in cortically related intermuscular coherence accompanying improvements in locomotor skills in incomplete spinal cord injury. *Journal of Neurophysiology, 95,* 2580-2589.

Ohala, J. Aspects of the control and production of speech. *Working Papers in Phonetics, 15,* 1-195.

Patel, R., & Schell, K.W. (2008). The influence of linguistic content on the Lombard effect. *Journal of Speech, Language, and Hearing Research, 51*(1), 209-220.

Perkell, J.S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics, 25,* 382-407.

Pick, H.L., Siegel, G.M., Fox, P.W., Garber, S.R., & Kearny, J.K. (1989). Inhibiting the Lombard effect. *Journal of the Acoustical Society of America, 85*(2), 894-900.

Rossiter, H.E., Eaves, C., Davis, E., Boudrias, M.H., Park, C., Farmer, S., Barnes, G., Litvak, V., & Ward, N.S. (2013). *NeuroImage: Clinical, 2,* 50-55.

Russell, N.K., & Stathopoulos, E. (1988). Lung volume changes in children and adults during speech production. *Journal of Speech, Language & Hearing Research, 31*(2), 146-155.

Sadagopan, N., Huber, J.E. (2007). Effects of loudness cues on respiration in individuals with Parkinson's disease. *Movement Disorders, 5,* 651-659.

Schneider, P., Dubé, R. V., & Hayward, D. (2005). The Edmonton Narrative Norms Instrument. Retrieved February 2 2016 from University of Alberta Faculty of Rehabilitation Medicine website: `http://www.rehabmed.ualberta.ca/spa/enni.`

Solomon, N.P., & Charron, S. (1998). Speech breathing in able-bodied children and children with cerebral palsy: A review of the literature. *American Journal of Speech-Language Pathology, 7*(2), 61-78.

Siegel, G.M., Pick, H.L., Olsen, M.G., & Sawin, L. (1976). Auditory feedback in the regulation of vocal intensity of preschool children. *Developmental Psychology*, *12*(3). 255-261.

Stathopoulos, E.T., Huber, J.E., Richardson, K., Kamphaus, J., DeCicco, D., Darling, M., Fulcher, K., Sussman, J.E. (2014). Increased vocal intensity due to the Lombard effect in speakers with Parkinson's disease: Simultaneous laryngeal and respiratory strategies. *Journal of Communication Disorders, 48,* 1-17.

Stathopoulous, E.T., & Sapienza, C.M. (1997). Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *Journal of Speech, Language & Hearing Research, 40* (3), 595-614.

Stepp, C. E., Hillman, R.E., & Heaton, J.T. (2010). Modulation of neck intermuscular beta coherence during voice and speech production. *Journal of Speech, Language, and Hearing Research, 54,* 836-844.

Smith, A., & Denny, M. (1990). High frequency oscillations as indicators of neural control mechanisms in human respiration, mastication, and speech. *Journal of Neurophysiology, 63*(4), 745-759.

Smitheran, J.R., & Hixon, T.J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders, 46*(2), 138-146.

Stowe, L.M., & Golob, E.J. (2013). Evidence that the Lombard effect is frequency-specific in humans. *Journal of the Acoustical Society of America, 134*(1), 640-647.

Tam, A., Cummine, J., Hodgetts, W., Tucker, B., & Boliek, C.A. (2016). The effect of attenuated acoustic feedback on vocal loudness in adults. Poster presented at the Faculty of Rehabilitation Medicine Research Day, University of Alberta, Edmonton, AB.

Tang, J., & Stathopoulos, E.T. (1995). Vocal efficiency as a function of vocal intensity: A study of children, women, and men. *Journal of the Acoustical Society of America, 97*(3), 1885-1892.

Ternström, S., Bohman, M., Södersten, M. (2006). Loud speech over noise: Some spectral attributes, with gender differences. *Journal of the Acoustical Society of America, 119*(3), 1648-1665.

Titze, I.R., (1988). On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the the Acoustical Society of America, 85,* 901-906.

Tomczak, C.R., Greidanus, K.R., & Boliek, C.A. (2013). Modulation of chest wall intermuscular coherence: effects of lung volume excursion and transcranial direct current stimulation. *Joural of Neurophysiology, 110,* 680-687.

Tufts, J.B., & Frank, &. (2003). Speech production in noise with and without hearing protection. *Journal of the Acoustical Society of America, 114*(2), 1069-1080.

Van Summers, W., Pisoni, D.B., Bernacki, R.H., Pedlow, R.I., Stokes, M.A. (1988). Effects of noise on speech production: Acoustical and perceptual analyses. *Journal of the Acoustical Society of America, 84*(3). 917-928.

Wang, Y-T., Green, J.R., Nip, I.S.B., Kent, R.D., & Kent, J.F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatrica et Logopaedica, 62*(6). 297-302.

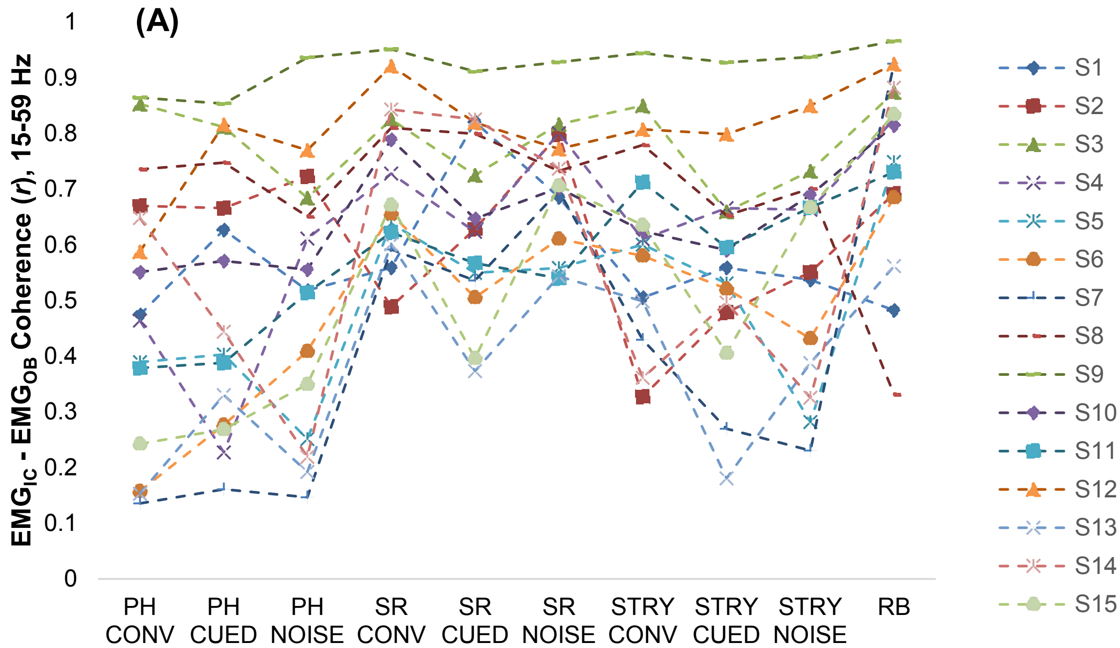
Wanger, J., Clausen, J.L., Coates, A., Pedersen, O.F., Brusasco, V., Burgos, F., . . . Viegi, G. (2005). Standardisation of the measurement of lung volumes. *European Respiratory Journal, 26,* 511-522.

Winkworth, A.L., Davis, P. (1997). Speech breathing and the Lombard effect. *Journal of Speech, Language, and Hearing Research, 40*(1), 159-169.

Winkworth, A.L., Davis, P.J., Ellis, E., & Adams, R.D. (1994). Variability and consistency in speech breathing during reading: lung volumes, speech intensity, and linguistic factors. *Journal of Speech, Language & Hearing Research, 37*(3), 535-556.

Winkworth, A.L., Davis, P.J., Adams, R.D., & Ellis, E. (1995). Breathing patterns during spontaneous speech. *Journal of Speech, Language & Hearing Research, 38*(1). 124-144.

Appendix A



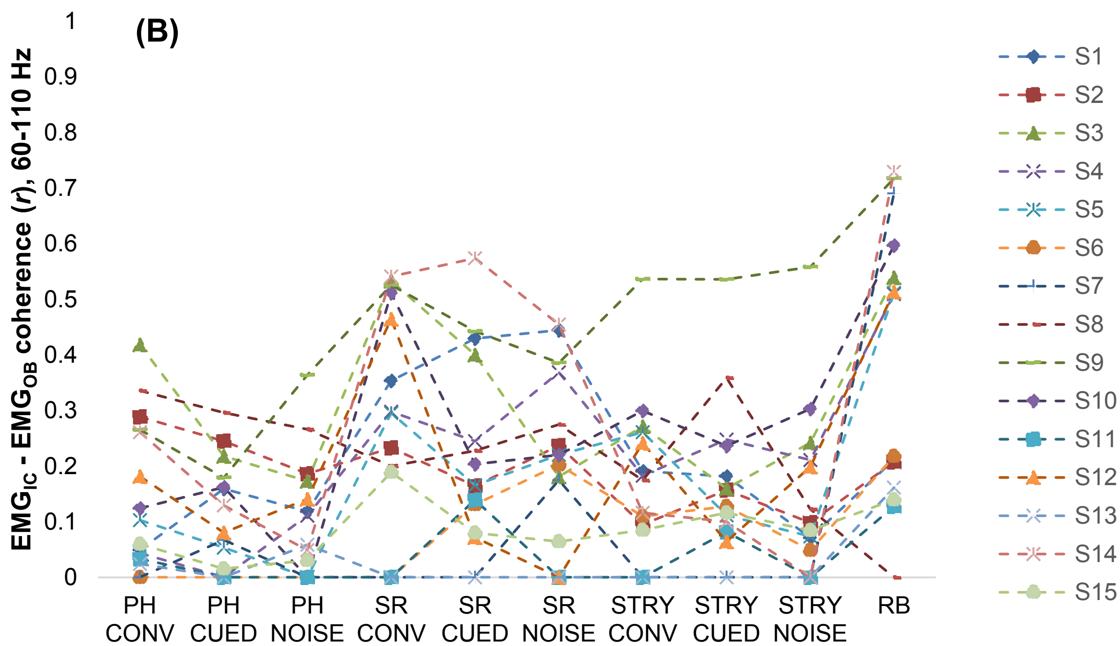


Fig 10. Mean peak intercostal-oblique intermuscular coherence values (*r*) for each subject (S1-15) in the low (A) and high (B) frequency bands, showing individual responsiveness to cue condition for each task. PH = phonation, SR = sentence repetition, STRY = storytelling, RB = resting breathing.